

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2003, 10:13:20 ; Search time 28 Seconds
(without alignments)
205.960 Million cell updates/sec

Title: US-09-831-758A-8
Perfect score: 1018
Sequence: 1 MEITSSKLFILLTLATSSLL.....KQSRRLFFKKIDDAELKQEK 196

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pap.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pap.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pap.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pap.*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pap.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76.5	7.5	530	PCT-US95-08493-2	Sequence 2, Appli
2	76.5	7.5	860	PCT-US95-08493-19	Sequence 19, Appli
3	76.5	7.5	868	PCT-US95-08493-21	Sequence 21, Appli
4	75.5	7.4	869	US-08-374-834-16	Sequence 16, Appli
5	75.5	7.4	869	US-08-644-271-29	Sequence 29, Appli
6	75.5	7.4	869	US-09-077-955-33	Sequence 33, Appli
7	75	7.4	725	US-09-594-669-16	Sequence 16, Appli
8	74.5	7.3	1170	US-09-749-588-2	Sequence 2, Appli
9	72.5	7.1	348	US-08-933-750C-10	Sequence 10, Appli
10	72.5	7.1	348	US-09-234-613-10	Sequence 10, Appli
11	72.5	7.1	723	US-09-594-669-14	Sequence 14, Appli
12	72	7.1	538	US-09-457-040B-23	Sequence 23, Appli
13	72	7.1	816	US-08-038-760-3	Sequence 3, Appli
14	72	7.1	816	US-08-470-091-3	Sequence 3, Appli
15	72	7.1	928	US-08-204-329-1	Sequence 1, Appli
16	72	7.1	928	US-08-959-638-8	Sequence 8, Appli
17	72	7.1	928	US-08-482-627-5	Sequence 5, Appli
18	72	7.1	928	US-08-801-092-4	Sequence 4, Appli
19	72	7.1	928	US-08-328-673A-8	Sequence 8, Appli
20	72	7.1	928	US-09-315-113-4	Sequence 4, Appli
21	72	7.1	928	PCT-US94-10357-2	Sequence 2, Appli
22	72	7.1	2210	US-09-309-572-7	Sequence 7, Appli
23	71	7.0	282	US-09-134-001C-3155	Sequence 3155, Ap
24	71	7.0	377	US-09-420-211-2	Sequence 2, Appli
25	71	7.0	928	PCT-US94-10357-3	Sequence 3, Appli
26	69.5	6.8	628	US-08-257-073-9	Sequence 9, Appli
27	69.5	6.8	2938	PCT-US94-00198-3	Sequence 3, Appli

28	68.5	6.7	793	4	US-09-523-849-32	Sequence 32, Appli
29	68.5	6.7	878	4	US-09-134-001C-4378	Sequence 4378, Ap
30	68	6.7	332	4	US-09-299-041-2	Sequence 2, Appli
31	68	6.7	332	4	US-09-299-041-4	Sequence 4, Appli
32	68	6.7	332	4	US-09-299-041-6	Sequence 6, Appli
33	68	6.7	539	2	US-08-978-182-3	Sequence 3, Appli
34	68	6.7	539	2	US-09-205-681-3	Sequence 3, Appli
35	68	6.7	2052	3	US-09-045-201A-2	Sequence 2, Appli
36	68	6.7	2052	4	US-09-619-062-2	Sequence 2, Appli
37	67.5	6.6	335	2	US-08-903-800A-6	Sequence 6, Appli
38	67.5	6.6	420	4	US-09-213-053-4	Sequence 4, Appli
39	67.5	6.6	1394	4	US-09-213-053-2	Sequence 2, Appli
40	67	6.6	1036	4	US-08-891-640-3	Sequence 3, Appli
41	67	6.6	1041	3	US-08-898-978-2	Sequence 2, Appli
42	67	6.6	1041	4	US-09-372-858-2	Sequence 2, Appli
43	67	6.6	1061	4	US-08-701-154A-5	Sequence 5, Appli
44	66.5	6.5	301	4	US-09-134-001C-3518	Sequence 3518, Ap
45	66.5	6.5	857	1	US-08-220-0151-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1
PCT-US95-08493-2
; Sequence 2, Application PC/TUS9508493
; GENERAL INFORMATION:
; APPLICANT: Wood, Clive
; APPLICANT: Caruso, Anthony
; TITLE OF INVENTION: Novel mlk Receptor Tyrosine Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEGAL AFFAIRS
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08493
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A
; REGISTRATION NUMBER: 32,724
; REFERENCE/DOCKET NUMBER: G15234A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-08493-2

Query Match 7.5%; Score 76.5; DB 5; Length 530;
Best Local Similarity 21.4%; Pred. No. 1.7;
Matches 43; Conservative 32; Mismatches 63; Indels 63; Gaps 13;

OY 1 MEIIS-KLFIILLTLATSSLLTSNIFCA-----DELVMNLHS 37
Db 159 ISIVSFFALLTLAT-----LYCCRRRREKWNKKREKSTAVTLTLPSELLDLRHP 211
OY 38 KENVDKYS---EPR---GYPKGRSLNFEELKDMGP---KNVIMSTPAVNKMPHS-FA 86
Db 212 NPMYQRMPLLLNPKLLSLLEYPRN-----NIEYVRDIGEGAFGRVQARAPGL--LPYEP 265

QY 87 NLPLRFRGNVOERSAGATANLPLSRGRNMEVSLVRRV--PNLPORFGRTRTTAKSVCRML 144
Db 266 MVAVKM---LKEASADMQADF-----OREAALMAEFDNPVIVKLLGVCAGVKPMCLLF 316
QY 145 SDCQSGMHSPCANDLFYSMT 165
Db 317 EYMAYGDL-----NEFLRSM 332

RESULT 2

PCT-US95-08493-19
; Sequence 19, Application PC/TUS9508493
; GENERAL INFORMATION:
; APPLICANT: Wood, Clive
; APPLICANT: Caruso, Anthony
; TITLE OF INVENTION: Novel mlk Receptor Tyrosine Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEGAL AFFAIRS
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08493
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A
; REGISTRATION NUMBER: 32,724
; REFERENCE/DOCKET NUMBER: G15234A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 860 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-08493-19

Query Match 7.5%; Score 76.5; DB 5; Length 860;
Best Local Similarity 21.4%; Pred. No. 3.4;
Matches 43; Conservative 32; Mismatches 63; Indels 63; Gaps 13;

QY 1 MEISS-KLFIILLTLATSLTSLTSLNIFCA-----LYCRRRKWKKKKREAVTLTLPSELLDLRLHP 541
Db 489 ISIVSSFALFALLTAT-----LYCRRRKWKKKKREAVTLTLPSELLDLRLHP 541
QY 38 KENYDKYS---EPR---GYPKGERSLNFEELKDWGP---KNVLMSTPAVNMKPHS-FA 86
Db 542 NPMYQRMPLLLNPKLLSLEYPRN---NIEYVRDIGEGAFGRVQARAPGL--LPYEPFT 595
QY 87 NLPLRFRGNVOERSAGATANLPLSRGRNMEVSLVRRV--PNLPORFGRTRTTAKSVCRML 144
Db 596 MVAVKM---LKEASADMQADF-----OREAALMAEFDNPVIVKLLGVCAGVKPMCLLF 646
QY 145 SDCQSGMHSPCANDLFYSMT 165
Db 647 EYMAYGDL-----NEFLRSM 662

RESULT 3

PCT-US95-08493-21
; Sequence 21, Application PC/TUS9508493

; GENERAL INFORMATION:
; APPLICANT: Wood, Clive
; APPLICANT: Caruso, Anthony
; TITLE OF INVENTION: Novel mlk Receptor Tyrosine Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEGAL AFFAIRS
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08493
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A
; REGISTRATION NUMBER: 32,724
; REFERENCE/DOCKET NUMBER: G15234A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 868 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-08493-21

Query Match 7.5%; Score 76.5; DB 5; Length 868;
Best Local Similarity 21.4%; Pred. No. 3.5;
Matches 43; Conservative 32; Mismatches 63; Indels 63; Gaps 13;
QY 1 MEISS-KLFIILLTLATSLTSLTSLNIFCA-----LYCRRRKWKKKKREAVTLTLPSELLDLRLHP 549
Db 497 ISIVSSFALFALLTAT-----LYCRRRKWKKKKREAVTLTLPSELLDLRLHP 549
QY 38 KENYDKYS---EPR---GYPKGERSLNFEELKDWGP---KNVLMSTPAVNMKPHS-FA 86
Db 550 NPMYQRMPLLLNPKLLSLEYPRN---NIEYVRDIGEGAFGRVQARAPGL--LPYEPFT 603
QY 87 NLPLRFRGNVOERSAGATANLPLSRGRNMEVSLVRRV--PNLPORFGRTRTTAKSVCRML 144
Db 604 MVAVKM---LKEASADMQADF-----OREAALMAEFDNPVIVKLLGVCAGVKPMCLLF 654
QY 145 SDCQSGMHSPCANDLFYSMT 165
Db 655 EYMAYGDL-----NEFLRSM 670

RESULT 4

US-08-374-834-16
; Sequence 16, Application US/08374834
; Patent No. 5656473
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, et al.
; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/374,834
; FILING DATE: 19-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/095,658
; FILING DATE: 21-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Cobert, Robert J.
; REGISTRATION NUMBER: 36,108
; REFERENCE/DOCKET NUMBER: REG 190A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 345-7400
; TELEFAX: (914) 345-7721
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 869 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-374-834-16

```

```

Query Match 7.48; Score 75.5; DB 1; Length 869;
Best Local Similarity 20.18; Pred. No. 4.6;
Matches 44; Conservative 37; Mismatches 69; Indels 69; Gaps 14;

QY 1 MEIIS-KLFILLTATSLTSNIFCA-----LYCRRRKQKKNKRESAAVTLTLPSELLDRLHP 550
Db 498 ISIMSSFAIVLLTIT-----LYCRRRKQKKNKRESAAVTLTLPSELLDRLHP 550
QY 38 KENYDKYS---EPR---GYPKGERSLNFEELKDWGP---KNVTKMSTPAVNKMPHS-FA 86
Db 551 NPMYQRMPLLNPKLLSLEYPRN---NIEVVRDIGEGAFGRVFOARAPGL--LPYEPFT 604
QY 87 NLPLRFGRNVOEERSAGATANLPLRSGRNWEVSLVRV--PNLPQRFGRTTTAKSVCRML 144
Db 605 MVAVKM---LKEASADMQADF-----QREAAALMAEFDNPNIKLLGVCAGVKPMCLLF 655
QY 145 SDLCQCG-----SMHSPCANDLFYSMTQHQEIQNP 174
Db 656 EYNMAYGDLNEFLRSMSPHTVCS--LSHSDLSMRAQVSSP 692

```

```

RESULT 5
US-08-644-271-29
; Sequence 29, Application US/08644271
; Patent No. 5814478
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, et al.
; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS
; NUMBER OF INVENTION: AND LIGANDS
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill Road
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/644,271
; FILING DATE: 10-MAY-1996
; CLASSIFICATION: 435

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USN 60/008,657
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Cobert, Robert J.
; REGISTRATION NUMBER: 36,108
; REFERENCE/DOCKET NUMBER: REG 195A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 914-345-7400
; TELEFAX: 914-345-7721
; TELEX:
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 869 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-644-271-29

```

```

Query Match 7.48; Score 75.5; DB 2; Length 869;
Best Local Similarity 20.18; Pred. No. 4.6;
Matches 44; Conservative 37; Mismatches 69; Indels 69; Gaps 14;

QY 1 MEIIS-KLFILLTATSLTSNIFCA-----LYCRRRKQKKNKRESAAVTLTLPSELLDRLHP 550
Db 498 ISIMSSFAIVLLTIT-----LYCRRRKQKKNKRESAAVTLTLPSELLDRLHP 550
QY 38 KENYDKYS---EPR---GYPKGERSLNFEELKDWGP---KNVTKMSTPAVNKMPHS-FA 86
Db 551 NPMYQRMPLLNPKLLSLEYPRN---NIEVVRDIGEGAFGRVFOARAPGL--LPYEPFT 604
QY 87 NLPLRFGRNVOEERSAGATANLPLRSGRNWEVSLVRV--PNLPQRFGRTTTAKSVCRML 144
Db 605 MVAVKM---LKEASADMQADF-----QREAAALMAEFDNPNIKLLGVCAGVKPMCLLF 655
QY 145 SDLCQCG-----SMHSPCANDLFYSMTQHQEIQNP 174
Db 656 EYNMAYGDLNEFLRSMSPHTVCS--LSHSDLSMRAQVSSP 692

```

```

RESULT 6
US-09-077-955-33
; Sequence 33, Application US/09077955A
; Patent No. 6413740
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al., David M.
; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS AND LIGANDS
; FILE REFERENCE: REG195-B-PCT-US
; CURRENT APPLICATION NUMBER: US/09/077,955A
; EARLIER FILING DATE: 1998-09-10
; EARLIER FILING DATE: 1996-12-13
; EARLIER FILING DATE: 1996-05-10
; EARLIER FILING DATE: 1995-12-15
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33
; LENGTH: 869
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-077-955-33

```

```

Query Match 7.48; Score 75.5; DB 4; Length 869;
Best Local Similarity 20.18; Pred. No. 4.6;
Matches 44; Conservative 37; Mismatches 69; Indels 69; Gaps 14;

QY 1 MEIIS-KLFILLTATSLTSNIFCA-----LYCRRRKQKKNKRESAAVTLTLPSELLDRLHP 550
Db 498 ISIMSSFAIVLLTIT-----LYCRRRKQKKNKRESAAVTLTLPSELLDRLHP 550

```

Qy	38	KENYDKYS	---EPR---	GYPKGERSINFEELKDWGP	---KNVTKMSTPAVKNMPSH-FA	86
			! :	! :	! :	
Db			! :	! :	! :	
	551	NPMYQRMPELLNPKLLSLEYPRN	---NIEYRDLGEGAFGRVQARAPGL	--LPYEPFT	604	
Qy	87	NLPLRFGNRVOERSAGATANLPLRSGRNMEYSLVRRY	--PNLPQRFQRTTAKSVCRM	144		
		! :	! :	! :	! :	
Db	605	MYAVKM	--LKEEASADMQADF	---QREAAALMAEFDNPINVKLLGVCAGKPMCLLF	655	
		! :	! :	! :	! :	
Qy	145	SDLCOG	-----SMHSPCANDLFYSMTCOHQEIQNP	174		
		! :	! :	! :	! :	
Db	656	EYMAVGDLNEFLRMSPHYTCS	--LSHSDLSMRAQVSSP	692		
		! :	! :	! :	! :	

TYPE: amino acid

RES-08-038-760-3
Sequence 3, Application US/08038760
Patent No. 5496731
GENERAL INFORMATION:
APPLICANT: Xu, Hong-Ji
APPLICANT: Hu, Shi-Xue
APPLICANT: Benedict, William F.
TITLE OF INVENTION: Broad-Spectrum Tumor Suppressors
TITLE OF INVENTION: Methods for Tumor Suppressor
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/038,760
FILING DATE: 19930325
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 7409-025-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-470-091-3

Query Match 7.1%; Score 72; DB 2; Length 816;
Best Local Similarity 21.5%; Pred. No. 11;
Matches 47; Conservative 23; Mismatches 63; Indels 86; Gaps 8;

QY 17 SSLTSNIF-----CADELVMS-----NLHSENYDKYSEPR 48
Db 362 SKLLNDNIFHMSLLACALEVVMATYSRSTSONLDSGTDLSFPWILNVLNKAFFDYKVIE 421
QY 49 GYPKGRSLNFEELK-----DWGPKNVIKMSTPAVNKMPHSFANLPLR 91
Db 422 SFIAKAGNLTREMIKHLERCEHRIMESLAWLSDSLPLDLIKQSKDREGTDLHESACPL- 480
QY 92 FGRLVQEERSAGATANLPLRSR-----NMEVSLVRRVNPQLPQRFGRITTTAKSVCRMISDLCL 148
Db 481 -----NLPLQNNHTAADMYLSPVRS-----PKKKGSTTRVNSTANAETQ-- 519
QY 149 QGSMHSPCANDLFYSMTQHOEIQNDQKOSRRLFPKKI 187
Db 520 -----ATSAF-----QTQKPLKSTLSLFFYKKV 542

RESULT 15

US-08-204-329-1
Sequence 1, Application US/08204329
Patent No. 5710255
GENERAL INFORMATION:
APPLICANT: SHEPARD, H. M.
TITLE OF INVENTION: CHARACTERIZATION OF A NOVEL ANTI-P110RB
TITLE OF INVENTION: MONOCLONAL ANTIBODY
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND & TOWNSEND & CREW LLP
STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: U.S.A.
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,329
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/05866
FILING DATE: 14-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: RENEE A. FITTS
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 16930-00040005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 928 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-204-329-1

Query Match 7.1%; Score 72; DB 1; Length 928;
Best Local Similarity 21.5%; Pred. No. 13;
Matches 47; Conservative 23; Mismatches 63; Indels 86; Gaps 8;

QY 17 SSLTSNIF-----CADELVMS-----NLHSENYDKYSEPR 48
Db 474 SKLLNDNIFHMSLLACALEVVMATYSRSTSONLDSGTDLSFPWILNVLNKAFFDYKVIE 533
QY 49 GYPKGRSLNFEELK-----DWGPKNVIKMSTPAVNKMPHSFANLPLR 91
Db 534 SFIAKAGNLTREMIKHLERCEHRIMESLAWLSDSLPLDLIKQSKDREGTDLHESACPL- 592
QY 92 FGRLVQEERSAGATANLPLRSR-----NMEVSLVRRVNPQLPQRFGRITTTAKSVCRMISDLCL 148
Db 593 -----NLPLQNNHTAADMYLSPVRS-----PKKKGSTTRVNSTANAETQ-- 631
QY 149 QGSMHSPCANDLFYSMTQHOEIQNDQKOSRRLFPKKI 187
Db 632 -----ATSAF-----QTQKPLKSTLSLFFYKKV 654

Search completed: March 26, 2003, 10:17:28
Job time : 32 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2003, 10:14:34 ; Search time 28 Seconds
(without alignments)
923.935 Million cell updates/sec

Title: US-09-831-758a-8

Perfect score: 1018

Sequence: 1 MEIISKFLTLATSSLL.....KQSRRLFKIDDAELKQEK 196

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 613006 seqs, 131990659 residues

Total number of hits satisfying chosen parameters: 613006

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Pending Patents_AA_New.*
- 1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
 - 2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
 - 3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
 - 4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
 - 5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
 - 6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
 - 7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	9.5	931	6	US-10-282-122A-56549
2	84	8.3	629	6	Sequence 56549, A
3	83.5	8.2	1817	6	Sequence 2002, Ap
4	81.5	8.0	1358	6	Sequence 3988, Ap
5	81.5	8.0	1426	6	Sequence 1658, Ap
6	81.5	8.0	1435	6	Sequence 15, Appl
7	81	8.0	471	5	Sequence 2178, Ap
8	81	8.0	471	5	Sequence 74897, A
9	81	8.0	502	5	Sequence 74898, A
10	81	8.0	537	5	Sequence 74898, A
11	81	8.0	537	5	Sequence 2042, Ap
12	81	8.0	709	5	Sequence 74897, A
13	81	8.0	709	5	Sequence 74897, A
14	81	8.0	713	5	Sequence 74897, A
15	81	8.0	713	5	Sequence 74897, A
16	81	8.0	713	5	Sequence 74897, A
17	81	8.0	713	5	Sequence 74897, A
18	81	8.0	713	5	Sequence 74897, A
19	81	8.0	713	5	Sequence 74897, A
20	80	7.9	302	6	Sequence 74897, A
21	80	7.9	302	6	Sequence 74897, A
22	80	7.9	302	6	Sequence 74897, A
23	80	7.9	302	6	Sequence 74897, A
24	80	7.9	302	6	Sequence 74897, A
25	79.5	7.8	509	6	Sequence 74897, A
26	79.5	7.8	1275	6	Sequence 22705, A

27 79 7.8 647 6 US-10-282-122A-71843
28 78.5 7.7 588 6 US-10-264-213-167
29 78.5 7.7 660 6 US-10-264-213-138
30 78.5 7.7 662 6 US-10-264-213-232
31 78.5 7.7 709 6 US-10-264-213-155
32 77 7.6 270 6 US-10-282-122A-60255
33 77 7.6 1131 6 US-10-366-683-31247
34 76.5 7.5 1357 6 US-10-369-493-5432
35 76 7.5 899 6 US-10-369-493-4010
36 75.5 7.4 460 5 US-09-724-676-82376
37 75.5 7.4 460 5 US-09-724-676-82376
38 75 7.4 725 5 US-09-849-602-20
39 74.5 7.3 280 1 PCT-US03-03161-2
40 74.5 7.3 280 1 PCT-US03-03161-2
41 74.5 7.3 490 1 PCT-US03-03161-4
42 74.5 7.3 490 6 US-10-350-385-4
43 74.5 7.3 492 1 PCT-US03-03161-7
44 74.5 7.3 492 6 US-10-350-385-7
45 74.5 7.3 538 5 US-09-724-676-74895

ALIGNMENTS

RESULT 1

US-10-282-122A-56549
; Sequence 56549, Application US/10282122A

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A
; CURRENT FILING DATE: 2003-02-20

; PRIOR FILING DATE: 2000-03-21

; PRIOR FILING DATE: 2000-05-23

; PRIOR FILING DATE: 2000-05-26

; PRIOR FILING DATE: 2000-09-06

; PRIOR FILING DATE: 2000-09-09

; PRIOR FILING DATE: 2000-10-23

; PRIOR FILING DATE: 2000-11-27

; PRIOR FILING DATE: 2000-12-22

; PRIOR FILING DATE: 2001-02-09

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 56549

; LENGTH: 931

; TYPE: PRT

; ORGANISM: Escherichia coli

; US-10-282-122A-56549


```

; PRIOR FILING DATE: 1999-04-02
; PRIOR APPLICATION NUMBER: 60/127,607
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 6322
; SOFTWARE: Curation Version 1.0
; SEQ ID NO 3988
; LENGTH: 1817
; TYPE: PRT
; ORGANISM: Homo sapiens
US -10-218-140-3988

Query Match      8.2%   Score 83.5; DB 6; Length 1817;
Best Local Similarity 21.2%; Pred No. 34;
Matches 35; Conservative 27; Mismatches 64; Indels 39; Gaps 5;

QY 28 DELVMSNIHSEKNDYKSEP-----RGYPKGRSLNFELKDWGPKNVTKM 73
   :||:|:|:| :||| ||| |||
Db 1527 EDLVSLDISPGKGVMDKILEPVACVRKKSEMIQLFPAYLKG-----DLFGL 1572

QY 74 STPAVKNPHSF-----ANLPLRFGNRVQERSAGATANLPLRSGRNMVSLVRRVPNL 127
   :||:|:|:| :||| ||| |||
Db 1573 TVSAVIAIESLPGVEACENTFYRIGLNLMELPLAVNPTGCARSEPKMSAH-VKRFVLR 1631

QY 128 PQRFGRTTAKSVCRMISDLCOGSMHSPCANDLFYSMTCOHOEIQ 172
   | :||:|:| :||| ||| |||
Db 1632 PHTLNSITSKS-----FQSTVTGELNAPYSKQFVHSKSSQYRKMK 1672

RESULT 4
US-10-369-493-1658
; Sequence 1658, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1658
; LENGTH: 1358
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-1658

Query Match      8.0%   Score 81.5; DB 6; Length 1358;
Best Local Similarity 23.2%; Pred. No. 39;
Matches 48; Conservative 34; Mismatches 84; Indels 41; Gaps 9;

QY 7 KLFILLTATSSLTNSIFCDELVMSNL-----HSEKNDYKSEPRGYPKGRSLNFEEL 62
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 880 RLQVLPYAVSTPENTNTEKNDVTNKNIGHQSEONSSSAKPSQIPTVSSPLGFEET 939

QY 63 KDWGPKNVTKMSTPAVKNPHSFANLPLRFGNRVQERSAGA-----TANLPLRSGRNM 117
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 940 K-----LSTPTKSNRNVSHSDTNSK--PKNTKENLKSWSRQEWLANLKL-----IS 986

QY 118 VSLVRRVPNLPQRFGRTTTAKSVCRMISDLCOGSMHSPCANDL-----FYSM 164
   ||||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 987 VSLVDFEPLSDDSQIINEXM-QLLKDIIFANNLKSAISNNPREDSDIILKGEIEDPYM 1045

QY 165 TCQ-----HOEION-PDOKOSRRLLFKK 186
   : : : : :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 1046 SSEIKIYYNELQNKPDAAKARFWSFMK 1072

RESULT 5

```


Matches 50; Conservative 44; Mismatches 79; Indels 64; Gaps 13;

```

QY 10 ILTLATSSLLTSNIP--CADELVMN--LHSKE-----NYDK-----YSEPRGYPKGER 55
DB 13 LTRMAMDSLSQARLPGLAIKIQRSNGLIHSANVRTVNLKSCVSEWAE--GGATKG-K 70
QY 56 SLNFEELKDWGPKNVKIMSTPAVNKMP--HSFANLPLRFRGNVVOERSAGATANLPL---- 110
DB 71 EIDFDDVAAINPE-----LLQLLPLHPKTNLPLQENVTIQKRRSVNSKIPAPKES 122
QY 111 ---RSGRNMESVLYR-----RVNPLPQRFGRTRTTAKSVCRMLS 145
DB 123 LRSRSTRMSTVSELRTAQENDMEVELPAAANSRKQFSVPPATPSPCPAFAEIPLRMVS 182
QY 146 DLCOGSMH-----SPCANDL--FYSMTCOHQEIQNPQKQSRRLFLFKKIDDAELKQEK 196
DB 183 EEMEEQVHSIRGSSSANPVNSVRRKSCLVKEVKMKNKREE----KKAQNSEMRMRK 235

```

RESULT 9

US-10-017-161-2042

; Sequence 2042, Application US/10017161

; GENERAL INFORMATION:

; APPLICANT: SUWA, MAKIKO

; APPLICANT: ASAI, KIYOSHI

; APPLICANT: AKIYAMA, YUTAKA

; APPLICANT: ABURATANI, HIROYUKI

; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS

; FILE REFERENCE: 084335/0152

; CURRENT APPLICATION NUMBER: US/10/017,161

; PRIOR FILING DATE: 2002-12-18

; PRIOR APPLICATION NUMBER: JP 2001/246789

; NUMBER OF SEQ ID NOS: 2430

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2042

; LENGTH: 502

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-017-161-2042

Query Match 8.0%; Score 81; DB 6; Length 502;
 Best Local Similarity 31.1%; Pred. No. 14;
 Matches 38; Conservative 15; Mismatches 49; Indels 20; Gaps 7;

```

QY 6 SKLFILLTATSSLLTSNIPFCADLVMSNL-----HSKENYDKYSEPRGYPKGER 55
DB 358 SGLF--LTVG--SSILTA--FLLSQLVQKLDQKTKETCRTHSRDRPRKLVGRGTAKGAR 413
QY 56 SLNFEELKDWGPKNVKIMSTPAVNKMP--HSFANLPLRFRGNVVOERSAGATANLPLRSGRN 115
DB 414 AL---PGRGWPRGRSRSPAPSVG--ARSERRR--RFRWSLRREPPAPASCRLLLSGHE 467
QY 116 ME 117
DB 468 AD 469

```

RESULT 10

US-09-724-676-74897

; Sequence 74897, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676

; NUMBER OF SEQ ID NOS: 2000-11-28

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 74897

; LENGTH: 537

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676-74897

Query Match 8.0%; Score 81; DB 5; Length 537;
 Best Local Similarity 21.1%; Pred. No. 16;
 Matches 50; Conservative 44; Mismatches 79; Indels 64; Gaps 13;

```

QY 10 ILTLATSSLLTSNIP--CADELVMN--LHSKE-----NYDK-----YSEPRGYPKGER 55
DB 13 LTRMAMDSLSQARLPGLAIKIQRSNGLIHSANVRTVNLKSCVSEWAE--GGATKG-K 70
QY 56 SLNFEELKDWGPKNVKIMSTPAVNKMP--HSFANLPLRFRGNVVOERSAGATANLPL---- 110
DB 71 EIDFDDVAAINPE-----LLQLLPLHPKTNLPLQENVTIQKRRSVNSKIPAPKES 122
QY 111 ---RSGRNMESVLYR-----RVNPLPQRFGRTRTTAKSVCRMLS 145
DB 123 LRSRSTRMSTVSELRTAQENDMEVELPAAANSRKQFSVPPATPSPCPAFAEIPLRMVS 182
QY 146 DLCOGSMH-----SPCANDL--FYSMTCOHQEIQNPQKQSRRLFLFKKIDDAELKQEK 196
DB 183 EEMEEQVHSIRGSSSANPVNSVRRKSCLVKEVKMKNKREE----KKAQNSEMRMRK 235

```

RESULT 11

US-09-724-676a-74897

; Sequence 74897, Application US/09724676a

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676a

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 74897

; LENGTH: 537

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676a-74897

Query Match 8.0%; Score 81; DB 5; Length 537;
 Best Local Similarity 21.1%; Pred. No. 16;
 Matches 50; Conservative 44; Mismatches 79; Indels 64; Gaps 13;

```

QY 10 ILTLATSSLLTSNIP--CADELVMN--LHSKE-----NYDK-----YSEPRGYPKGER 55
DB 13 LTRMAMDSLSQARLPGLAIKIQRSNGLIHSANVRTVNLKSCVSEWAE--GGATKG-K 70
QY 56 SLNFEELKDWGPKNVKIMSTPAVNKMP--HSFANLPLRFRGNVVOERSAGATANLPL---- 110
DB 71 EIDFDDVAAINPE-----LLQLLPLHPKTNLPLQENVTIQKRRSVNSKIPAPKES 122
QY 111 ---RSGRNMESVLYR-----RVNPLPQRFGRTRTTAKSVCRMLS 145
DB 123 LRSRSTRMSTVSELRTAQENDMEVELPAAANSRKQFSVPPATPSPCPAFAEIPLRMVS 182
QY 146 DLCOGSMH-----SPCANDL--FYSMTCOHQEIQNPQKQSRRLFLFKKIDDAELKQEK 196
DB 183 EEMEEQVHSIRGSSSANPVNSVRRKSCLVKEVKMKNKREE----KKAQNSEMRMRK 235

```

RESULT 12

US-09-724-676-74919

; Sequence 74919, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 74919

[illegible]

```

; LENGTH: 709
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-74919

Query Match      8.0%; Score 81; DB 5; Length 709;
Best Local Similarity 21.1%; Pred. No. 21;
Matches 50; Conservative 44; Mismatches 79; Indels 64; Gaps 13;

QY   10 ILLTATSSLLTSNIF--CADELVMNSN--LHSKE-----NYDK-----YSEPRGYPKGER 55
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB   13 LTLRWAMDSSLQAARLPGLAIRQSNGLIHSANVRTVNLEKSCSVSEWAE--GGATKG-K 70
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY   56 SLNFEEELKDWCQPNVIMKSTPAVNKKMP-HSFANPLRFGRNVQEERSAGATANLPL---- 110
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB   71 EIDFDVVAAINPE-----LLQLLPHLPKTNLPQENVTIQOKRRSVNSKIPAPKES 122
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY   111 ---RSGRNMEVSLVR-----RVNPLPQRFGRTTTAKSVCRMLS 145
       ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB   123 LRSRSTRMSTVSELRIETAQENDEVELPAAANSRKQFSVPAPTRPSCPVAEIPLRMVS 182
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY   146 DLCOGSMH-----SPCANDL--FYSMTCHOHQEIQNPDQOSRRLFLFKKIDDALCKOEK 196
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB   183 EEMEBQVHSIRGSSSANPVNSVRRKSCLVKEVKMKNKREE---KKAQNSEMRMCR 235
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 13
US-09-724-676A-74919
; Sequence 74919, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compuen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 74919
; LENGTH: 709
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-74919

Query Match      8.0%; Score 81; DB 5; Length 709;
Best Local Similarity 21.1%; Pred. No. 21;
Matches 50; Conservative 44; Mismatches 79; Indels 64; Gaps 13;

QY   10 ILLTATSSLLTSNIF--CADELVMNSN--LHSKE-----NYDK-----YSEPRGYPKGER 55
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB   13 LTLRWAMDSSLQAARLPGLAIRQSNGLIHSANVRTVNLEKSCSVSEWAE--GGATKG-K 70
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY   56 SLNFEEELKDWCQPNVIMKSTPAVNKKMP-HSFANPLRFGRNVQEERSAGATANLPL---- 110
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB   71 EIDFDVVAAINPE-----LLQLLPHLPKTNLPQENVTIQOKRRSVNSKIPAPKES 122
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY   111 ---RSGRNMEVSLVR-----RVNPLPQRFGRTTTAKSVCRMLS 145
       ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB   123 LRSRSTRMSTVSELRIETAQENDEVELPAAANSRKQFSVPAPTRPSCPVAEIPLRMVS 182
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY   146 DLCOGSMH-----SPCANDL--FYSMTCHOHQEIQNPDQOSRRLFLFKKIDDALCKOEK 196
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB   183 EEMEBQVHSIRGSSSANPVNSVRRKSCLVKEVKMKNKREE---KKAQNSEMRMCR 235
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 14
US-09-724-676-74922
; Sequence 74922, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compuen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28

```

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: March 26, 2003, 10:12:39 ; Search time 43 Seconds
(without alignments)
438.194 Million cell updates/sec

Title: US-09-831-758A-8
Perfect score: 1018
Sequence: 1 MEIISKLFIILLTATSSLL.....KQSRLLFKKIDDAELKQEK 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*

- 1: pir1.*
- 2: pir2.*
- 3: pir3.*
- 4: pir4.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	9.6	426	C96804	hypothetical prote
2	97	9.5	931	A64903	probable zinc prot
3	96	9.4	931	C90891	probable peptidase
4	96	9.4	931	F85726	probable peptidase
5	86.5	8.5	1211	D64116	exodeoxyribonuclea
6	85.5	8.4	1015	T32984	hypothetical prote
7	84	8.3	629	S63369	hypothetical prote
8	83.5	8.2	1163	J50366	tight junction pro
9	82	8.1	261	T25206	hypothetical prote
10	81.5	8.0	1358	A29360	SIR4 protein - yea
11	81	8.0	1343	A29360	cell division prot
12	80	7.9	1084	A196611	DNA-directed DNA p
13	80	7.9	1086	T43266	DNA-directed DNA p
14	80	7.9	1086	T40242	DNA polymerase del
15	80	7.9	1305	A80168	probable cell divi
16	79.5	7.8	144	G90573	50S ribosomal prot
17	79.5	7.8	327	D95408	probable oxidoredu
18	79.5	7.8	449	Q4ADA5	maturation (piva2)
19	79.5	7.8	449	Q4ADA2	maturation (piva2)
20	79.5	7.8	1275	T41523	hypothetical rhoI
21	79.5	7.8	1957	A45627	myosin heavy chain
22	79.5	7.8	609	S55957	hypothetical prote
23	78.5	7.7	688	T13248	NADH2 dehydrogenas
24	78	7.7	1204	T19918	hypothetical prote
25	77	7.6	448	Q4ADA7	maturation (piva2)
26	77	7.6	579	H87451	oxidoreductase, GM
27	77	7.6	1051	E83330	RND divalent metal
28	76.5	7.5	420	A37343	keratin 13, type I
29	76.5	7.5	458	KRHU3	keratin 13, type I

30	76.5	7.5	473	2	H84550	probable obtusifol
31	76.5	7.5	1357	2	T16860	hypothetical prote
32	76	7.5	1436	2	B81704	conserved hypothet
33	76	7.5	1893	2	T22661	hypothetical prote
34	75.5	7.4	361	2	T40857	hypothetical prote
35	75.5	7.4	371	2	T13023	drought-inducible
36	75.5	7.4	1642	2	T19130	hypothetical prote
37	75	7.4	517	2	D86637	ABC transporter AT
38	75	7.4	551	2	T03793	calmodulin-binding
39	75	7.4	582	2	H95950	probable ABC trans
40	74.5	7.3	492	2	T14820	obtusifolios 14-al
41	74.5	7.3	943	2	S28400	gag-like protein -
42	74.5	7.3	1073	2	T33764	hypothetical prote
43	74.5	7.3	1683	2	T30885	complement compone
44	74.5	7.3	1963	1	MWKW	myosin heavy chain
45	74	7.3	688	2	T13253	NADH2 dehydrogenas

ALIGNMENTS

RESULT 1

C96804
hypothetical protein T5M16.8 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 24-Aug-2001
C:Accession: C96804
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Jansen, N.F.; Hughes, B.; Huizlar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Rizzo, M.; Rooney, T.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Maier, M.; Salzman, S.L.; Rowley, D.; Sakano, H.; Schmitt, J.C.; Venter, J.C.; Davis, R.W.; Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: C96804
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-426 <STO>
A:Cross-references: GB:AE005173; NID:96382497; PIDN:AAF07783.1; GSPDB:GN00141
C:Genetics:
A:Gene: T5M16.8
A:Map position: 1
C:Superfamily: cytochrome-c peroxidase

Query Match	9.6%	Score 98;	DB 2;	Length 426;
Best Local Similarity	29.3%;	Pred. No. 0.16;		
Matches 53;	Conservative 20;	Mismatches 66;	Indels 42;	Gaps 11;
Qy	2	EIIISKPELLTATSSLLTSNIFCADELVNSLNHNSKENYDKYSEP---RCYPKGERSLN	58	
		::: :		

RESULT 2

A64903
probable zinc proteinase yddC (EC 3.4.99.-) - Escherichia coli (strain K-12)
C:Species: Escherichia coli

probable peptidase ppgL [imported] - Escherichia coli (strain O157:H7, substrain EDL_933)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: F85726
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May, D.A.;
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, N.;
Rabinovitch, A.; White, O.; Salzberg, S.L.; Peterson, J.; Fouts, D.E.; McIninch, J.;
Nelson, K.E.; Salzberg, S.L.; Smith, T.F.; Venter, A.; White, O.; Adams, M.D.;
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: F85726
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-931 <STO>
A:Cross-references: GB:AE005174; NID:q12515180; PIDN:AAG56274.1; GSPDB:GN00145; UWGFI:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: ppgL

	Query Match	9.4%	Score 96;	DB 2;	Length 931;	
	Best Local Similarity	27.3%;	Pred. No.	0.65;		
	Matches	38;	Conservative	27;	Mismatches	55; Indels 24; Gaps
Qy	1 MEIISSKLFIILLTATSSLLTSNIFCA-----DELVMSNLHSEKNDYKSYSPRGYPKGKER	55				
Dd	1 MEIIIMRNLCFLITLVATLLPGLRIAAALPODEKLITGOLDNGLRMYIY--PHAHPKDOV	58				
Qy	56 SLNF-----EELKDWGPRKNVIK-----MSTPAVNKMPSFANLPFRGRNNVOERS	101				
Dd	59 NLWLQIHGTSGLEEDNERGVAFHEVMFMFNCTKTWPNGKVITETESMCLRTGRDYNAYS	118				
Qy	102 AGAT---ANLPILSRGNME	117				
Dd	119 YDETIVQVSPLTTOKONLQ	137				

RESULT 5

D64116
exodeoxyribonuclease V 135K chain homolog - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999
C:Accession: D64116
R:Flaeschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavac,
J., D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vent
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7543800
A:Accession: D64116
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1211 <TIGR>
A:Cross-references: GB:U32811; GB:L42023; NID:g1574774; PIDN:AAC22966.1; PID:g157476
C:Superfamily: exodeoxyribonuclease V 135K chain

[illegible]

QY 166 CQHQEIQ 172
:| :::

Db 1070 REHFDVE 1076

RESULT 6

T32984

hypothetical protein K02D7.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T32984

R:Du, Z.; Magg1, L.

submitted to the EMBL Data Library, February 1998

A:Description: The sequence of C. elegans cosmid K02D7.

A:Reference number: Z21259

A:Accession: T32984

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1015 <DU2>

A:Cross-references: EMBL:AF045645; PIDN:AAC02607.1; GSPDB:GN00022; CESP:K02D7.4

A:Experimental source: strain Bristol N2; clone K02D7

C:Genetics:

A:Gene: CESP:K02D7.4

A:Map position: 4

A:Introns: 51/1; 103/1; 188/3; 304/2; 339/2; 392/3; 448/2; 545/3; 646/3; 794/3; 910/2

Query Match 8.4%; Score 85.5; DB 2; Length 1015;

Best Local Similarity 19.8%; Pred. No. 7.4; Indels 81; Gaps 8;

Matches 52; Conservative 35; Mismatches 94; Indels 81; Gaps 8;

QY 12 LTATSSLTSTNIFCADELVMSNLHSENYDKYSEPRGYPKGRSLNFEELKDWGPKNVI 71

Db 464 ITTAREVLLTASPDYLDLLFGISQSSSNNEKWKQLMYLWGLSLDKSEY--WKVANTI 521

QY 72 -----KMSTPAVN-----KMFHSFANLPL-----R 91

Db 522 ATVLNKRCEASTSSLNCKGKRTIVNKFITDLTAGGVEVRVLEVINIPIFGSTYFAKK 581

QY 92 FGQNVQERSAGATANLPLSRGRNNEVSLVRVNPQLPQRFGRRTTAKSVCRMLSLDLCQ-- 149

Db 582 FICETESEDVQKALNVILAASKNLYETQVSLKYEQEMRGYSAAHAEFFRHIS-ICQPK 640

QY 150 -----GSMHSPCANDLFYSWTCQHQ-----EIQNPQOK 177

Db 641 KKIGAPLTHKLIKLFRTNCSQETPTSHSOLAIDLKLCVPDQHVQNTLILRTLTLPDQO 700

QY 178 QSRRLFKKIDDA----ELKQE 195

Db 701 EKWHYLYKAIEASGNKDELKAE 722

RESULT 7

S63369

hypothetical protein YNR038w - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein N3302

C:Species: Saccharomyces cerevisiae

C:Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 02-Feb-2001

C:Accession: S63369

R:Pohl, T.M.

submitted to the Protein Sequence Database, April 1996

A:Reference number: S63346

A:Accession: S63369

A:Molecule type: DNA

A:Residues: 1-629 <POH>

A:Cross-references: EMBL:Z71653; NID:g1302540; PIDN:CAA96318.1; PID:e239589; PID:g130254

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:DBP6

A:Cross-references: SGD:S0005321; MIPS:YNR038w

A:Map position: 14R

C:Superfamily: fruit fly gene Dbp73D protein

C:Keywords: ATP; nucleotide binding; P-loop

F:234-241/Region: nucleotide-binding motif A (P-loop)

F:337-342/Region: nucleotide-binding motif B

F:341-344/Region: DEAD motif

Query Match 8.3%; Score 84; DB 2; Length 629;

Best Local Similarity 21.7%; Pred. No. 5.7; Indels 44; Gaps 6;

Matches 34; Conservative 28; Mismatches 51; Indels 44; Gaps 6;

QY 52 KGRSLNFEELKDWGPKNVIKMTSPAVNKPMPHSFANLPLRFRGNVQERSAGATANPLR 111

Db 342 EADRLN-QSFGQWCPKMLSHLTKDLDTP-----GNVIMFISATLTTNTEKL 390

QY 112 SGRNM-----EVSILVRVNPQLPQRFGR--RTTAKSVCR---MLSDLCGSMHSPCAND 159

Db 391 NGLNLYKPKLFLKQTDKLYQLPNKLNENINPTAKSVYKPLILLYLISICOFMAHSPIA-- 448

QY 160 LFYSMTCHQOEIQNPDKQSRRLIFKIDDAELKQEK 196

Db 449 -----AKILIFVKSNESSIRLSK 466

RESULT 8

JE0366

tight junction protein, ZO-2 - chicken

C:Species: Gallus gallus (chicken)

C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 21-Jul-2000

C:Accession: JE0366

R:Collins, J.R.; Rizzolo, L.J.

Biochem. Biophys. Res. Commun. 252, 617-622, 1998

A:Title: Protein-binding domains of the tight junction protein, ZO-2, are highly

A:Reference number: JE0366; MUID:99057550; PMID:9837755

A:Accession: JE0366

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1163 <COL>

A:Cross-references: GB:AF085184; NID:g3820579; PIDN:AAC95469.1; PID:g3820580

Query Match 8.2%; Score 83.5; DB 2; Length 1163;

Best Local Similarity 27.6%; Pred. No. 14; Indels 19; Gaps 3;

Matches 24; Conservative 17; Mismatches 27; Indels 19; Gaps 3;

QY 50 YPKGRSLNFEELKDWGPKN-----VIKMTSPAVNKPMPHSFANLPLRFRGNVQER 100

Db 34 FENGETSIVISDVLPGGPADGLLQENDRVIVNGTPEMEN-VPHSFA-----VQQLR 83

QY 101 SAGATANLPLSRGRNNEVSLVRVNPQL 127

Db 84 KSGKVATIVVKKRPKVQAAALRKNPSL 110

RESULT 9

T23206

hypothetical protein T23G7.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T23206

R:Barlow, K.

submitted to the EMBL Data Library, December 1995

A:Reference number: Z19995

A:Accession: T23206

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-261 <WIL>

A:Cross-references: EMBL:Z68319; PIDN:CAA92703.1; GSPDB:GN00020; CESP:T23G7.5

A:Experimental source: clone T23G7

C:Genetics:

A:Gene: CESP:T23G7.5

A:Map position: 2

A:Introns: 56/2; 137/2; 170/1; 207/1

Query Match 8.1%; Score 82; DB 2; Length 261;

Best Local Similarity 26.0%; Pred. No. 3; Indels 22; Gaps 7;

Matches 38; Conservative 23; Mismatches 63; Indels 22; Gaps 7;

QY 12 LTATSSLTSTNIFCADELVMSNLHSENYDKYSEPRGYPK--GERSLNFEELKDWGPKN 69

Db 13 LNLSHSLI--SIFSVS--VMSYHHNNHYQ--HRPRGYERLPKGLPDRWNIYDNGRD 66
QY 70 V-----IKMSTPAVNKMPHSFANLPLRFGRNVQERSAGATANLPLRSGRNMEVSLVRV 124
Db 67 IDGTRFVPPKTP---LDSSFFD-----CKNMPVELQFGVKTLISLAQAQANKQIGLVIDL 117
QY 125 PNLPRQFGRTRTTAKSVCRMLSDLCQG 150
Db 118 TNTDRIYKKTENADHGKVKLKLNCPG 143
RESULT 10
A29360
SIR4 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YD934.12; protein YDR227w; STE9 protein
C;Species: Saccharomyces cerevisiae
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 21-Jul-2000
C;Accession: A29360; S47935; S59434; S53988; S47476; S47950
R;Marshall, M.; Mahoney, D.; Rose, A.; Hicks, J.B.; Broach, J.R.
Mol. Cell. Biol. 7, 4441-4452, 1987
A;Title: Functional domains of SIR4, a gene required for position effect regulation in S
A;Reference number: A29360; MUID:88142836; PMID:3325825
A;Accession: A29360
A;Molecule type: DNA
A;Residues: 1-1358 <MAR>
A;Cross-references: GB:M37249; NID:g531115; PIDN:AAA20881.1; PID:g531116
R;Davies, C.J.; Hutchison III, C.A.
submitted to the EMBL Data Library, September 1994
A;Description: Tn3 transposon/deletion sequencing of a 9.4kb DNA fragment: Characterisat
A;Reference number: S47932
A;Accession: S47935
A;Molecule type: DNA
A;Residues: 1-993, 'L', 995-1358 <DAV>
A;Cross-references: EMBL:L35344; EMBL:U13239; EMBL:Z36548
R;Murphy, L.; Harris, D.
submitted to the EMBL Data Library, March 1995
A;Reference number: S59423
A;Accession: S59434
A;Molecule type: DNA
A;Residues: 1-1358 <MUR>
A;Cross-references: EMBL:Z48612; NID:g728671; PID:g728683; MIPS:YDR227w
A;Experimental source: strain AB972
R;Davies, C.J.; Hutchison III, C.A.
Nucleic Acids Res. 23, 507-514, 1995
A;Title: Insertion site specificity of the transposon Tn3.
A;Reference number: S53985; MUID:95192063; PMID:7885847
A;Accession: S53988
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-993, 'L', 995-1358 <DAW>
A;Cross-references: EMBL:U13239; NID:g532747; PIDN:ANC33144.1; PID:g532751
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994
C;Genetics:
A;Gene: SGD: SIR4; STE9
A;Cross-references: SGD:S0002635; MIPS:YDR227w
A;Map position: 4R
C;Keywords: DNA binding; transcription regulation
Query Match 8.0%; Score 81.5; DB 2; Length 1358;
Best Local Similarity 23.2%; Pred. No. 26;
Matches 48; Conservative 34; Mismatches 84; Indels 41; Gaps 9;
QY 7 KLFILLIATSLTSNIFCADELVMNL-----HSKENYDKYSEPRGYPKGRSLNFEEL 62
Db 880 RLQYLPYAVSTFNTNTEKNDVTNKNIGKHSQBNSSAKPSQIPTVSSPLGFEET 939
QY 63 KDWGPKNVKMTSPAVNKMPSFANLPLRFGRNVQERSAGA-----TANLPLRSGRNME 117
Db 940 K-----LSTPTKSNRVSHSDTNSK--PKNTKENLKSRSRQEWLANLKL-----IS 986
QY 118 VSLVRVNPPLRFGRTRTTAKSVCRMLSDLCQSGSMHSPCANDL-----FYSM 164
Db 987 VSLVDFPSELSDSQIINEKM-QLLKDIIFANLKLKSAISNNFRESDIILKGEIEDYPM 1045

QY 165 TCQ-----HOEIQN-PDQKOSRRLLFKK 186
Db 1046 SSEIKIYYNELQKPDAKKAREWSEMK 1072

RESULT 11

AF0611
cell division protein FtsK [imported] - Salmonella enterica subsp. enterica serovar
C;Species: Salmonella enterica subsp. enterica serovar typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C;Accession: AF0611
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churc
th, T.; Connerthon, P.; Cronin, A.; Davis, R.M.; Dowd, L.; White, N.; Fai
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Ouail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica ;
A;Reference number: AB0502; PMID:11677608
A;Accession: AF0611
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1343 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD05360.1; PID:g16502124; GSPDB:GN00176
C;Genetics:
A;Gene: STY0958

Query Match 8.0%; Score 81; DB 2; Length 1343;
Best Local Similarity 25.2%; Pred. No. 29;
Matches 29; Conservative 24; Mismatches 42; Indels 20; Gaps 6;
QY 65 WGPKNVVKMTSPAVNKMPSFANLPLRFGRNVQ-----SPANLPLRFGRNVQ-----ERSAGATANLPLRS 112
Db 1111 WKPQDSMDVQHPVLEKLPYIVLVVDEPADLMVTGKVEELIARLAKAARAAGTHLVAT 1170
QY 113 GR---NNEVSLVRVNPPLRFGRTRTTAKSVCRMLSDLCQSGSMHSPCA-NDLFVS 163
Db 1171 QRPSVDVITGLIK--ANIPTRIAFTVSSKIDSRITLD--OGGAESLLGMGMDLVS 1221

RESULT 12

S19661
DNA-directed DNA polymerase (EC 2.7.7.7) III large chain - fission yeast (Schizosac
C;Species: Schizosaccharomyces pombe
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S19661
R;Pignede, G.; Bouvier, D.; de Recondo, A.M.; Baldacchi, G.
J. Mol. Biol. 222, 209-218, 1991
A;Title: Characterization of the POL3 gene product from Schizosaccharomyces pombe fr
A;Reference number: S19661; MUID:92071954; PMID:1960723
A;Accession: S19661
A;Molecule type: DNA
A;Residues: 1-1084 <PIG>
A;Cross-references: EMBL:X59278; NID:g5010; PIDN:CAA41968.1; PID:g5011
C;Genetics:
A;Introns: 77/1
C;Superfamily: herpesvirus DNA-directed DNA polymerase
C;Keywords: DNA binding; nucleotidyltransferase; nucleus

Query Match 7.9%; Score 80; DB 1; Length 1084;
Best Local Similarity 27.9%; Pred. No. 27;
Matches 38; Conservative 20; Mismatches 36; Indels 42; Gaps 11;

QY 71 IKMSTPAVNKMPHSFANLPLRFGRNVQERSAGATANLPLRSGR-----NNEVSLVR 122
Db 971 ISMAAPSVGGI-----MKFA--VKVETCLGCKA--PIKKGKLTALCNCNLSAELYQ 1018
QY 123 R----VPLNPRQFGRTRTTAKSVCRMLSDLCQSGMHSP--CAN---DLFYSMTQCHOEIQN 173
Db 1019 ROVAQVNDLVRFARLWTO---COR----COGSMHQDVICTSRDPCIFYMRIAHHKLIQ- 1070
QY 174 PDQKOSRRLLFKKITDD 189

Db 1071 ----QSVDLL-KRFE 1081

RESULT 13
T43266
DNA-directed DNA polymerase (EC 2.7.7.7) delta chain - fission yeast (Schizosaccharomyces pombe)
N:Alternate names: DNA polymerase delta
C:Species: Schizosaccharomyces pombe
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C:Accession: T43266
R:Pignede, G.; Bouvier, D.; de Recondo, A.M.; Baldacci, G.
J. Mol. Biol. 222, 209-218, 1991
A:Title: Characterization of the POL3 gene product from Schizosaccharomyces pombe indica
A:Reference number: S19661; MUID:92071954; PMID:1960723
A:Accession: T43266
A:Molecule type: DNA
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Cross-references: EMBL:L07734; MID:g173383; PIDN:AAA35303.1; PID:g173384
C:Genetics:
A:Gene: poid
A:Introns: 77/1
C:Superfamily: herpesvirus DNA-directed DNA polymerase
C:Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 7.9%; Score 80; DB 2; Length 1086;
Best Local Similarity 27.9%; Pred. No. 27;
Matches 38; Conservative 20; Mismatches 36; Indels 42; Gaps 11;
QY 71 IKMSTPAVKNKPHSPANLPLRFRGNVQERSAGATANLPLRSGR-----NMEVSLVR 122
DB 973 ISMAAPSVGGI-----MKFA--VKVETCLGCKA--PIKKGKTAICLNCNLSAELYQ 1020
QY 123 R----VPNLPQRFGRITTAQSVKRMVSLDLCOGSMHSP--CAN---DLFYSMTQCQHOEQ 173
DB 1021 ROVAQVNDLEVRFAFLWTQ---COR-----COGSMHQDVICTSRDCPIFYMRIAEHKKLQ- 1072
QY 174 PDQKQSRRLFFKKIDD 189
DB 1073 ----QSVDDL-KRFE 1083

RESULT 14
T40242
DNA polymerase delta large chain - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C:Accession: T40242
R:Borzum, K.; Beck, A.; Reinhardt, R.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, October 1999
A:Reference number: Z21916
A:Accession: T40242
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1086 <BOR>
A:Cross-references: EMBL:AL121815; PIDN:CAB58156.1; GSPDB:GN00067; SPDB:SPBC336.04
A:Experimental source: strain 972h-; cosmid c336
C:Genetics:
A:Gene: SPDB:SPBC336.04
A:Map position: 2
A:Introns: 77/1
C:Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 7.9%; Score 80; DB 2; Length 1086;
Best Local Similarity 27.9%; Pred. No. 27;
Matches 38; Conservative 20; Mismatches 36; Indels 42; Gaps 11;
QY 71 IKMSTPAVKNKPHSPANLPLRFRGNVQERSAGATANLPLRSGR-----NMEVSLVR 122
DB 973 ISMAAPSVGGI-----MKFA--VKVETCLGCKA--PIKKGKTAICLNCNLSAELYQ 1020
QY 123 R----VPNLPQRFGRITTAQSVKRMVSLDLCOGSMHSP--CAN---DLFYSMTQCQHOEQ 173

Db 1021 ROVAQVNDLEVRFAFLWTQ---COR-----COGSMHQDVICTSRDCPIFYMRIAEHKKLQ- 1072
QY 174 PDQKQSRRLFFKKIDD 189
DB 1073 ----QSVDDL-KRFE 1083
RESULT 15
AB0168
probable cell division protein ftsK [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C:Accession: AB0168
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentic
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Douga
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Ba
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AB0168
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1305 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC90205.1; PID:g15979425; GSPDB:GN00175
C:Genetics:
A:Gene: ftsK
Query Match 7.9%; Score 80; DB 2; Length 1305;
Best Local Similarity 25.4%; Pred. No. 34;
Matches 29; Conservative 24; Mismatches 41; Indels 20; Gaps 6;
QY 65 WGPKNVKNKSTPAVKNKMPH-----SFANLPLRFRGNVQ-----ERSAGATANLPLRS 112
DB 1072 WKPSDSMDISPPMLVKLPYVVWVDFADLMTVGVKVEELIARLAQKARAAGIHLVLAT 1131
QY 113 GR---NMEVSLVRVNPVLPQRFGRITTAQSVKRMVSLDLCOGSMHSPCA--NDLFY 162
DB 1132 QRPSVDVITGLIK--ANIPTRIAFTVSSKIDSRITLD--QGAESLLGMDMLY 1181
Search completed: March 26, 2003, 10:16:54
Job time : 48 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2003, 10:08:54 ; Search time 26 Seconds
(without alignments)
312.668 Million cell updates/sec

Title: US-09-831-758A-8
Perfect score: 1018
Sequence: 1 MEITSSKLFILLTATSSLL.....KQSRRLFFKKIDDAELKQEK 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	97	9.5	931	1	POOL_ECOLI
2	86.5	8.5	1211	1	EX5B_HAEIN
3	84.5	8.3	685	1	SNWA_DICDI
4	84	8.3	629	1	DBP6_YEAST
5	81.5	8.0	1358	1	SIR4_YEAST
6	80	7.9	1086	1	DPOD_SCHPO
7	79.5	7.8	449	1	PIV2_ADE02
8	79.5	7.8	449	1	PIV2_ADE05
9	79	7.8	609	1	YL01_YEAST
10	77	7.6	448	1	PIV2_ADE07
11	76.5	7.5	446	1	PIV2_ADE40
12	76.5	7.5	458	1	K1CM_HUMAN
13	75	7.4	725	1	MCAC_HUMAN
14	74.5	7.3	492	1	CP51_SORBI
15	74.5	7.3	1966	1	MYSB_CAEEL
16	74	7.3	662	1	PDIL_MOUSE
17	73.5	7.2	643	1	NOS2_ALCEU
18	73.5	7.2	1131	1	AC15_MOUSE
19	73.5	7.2	1794	1	YAVI_SCHPO
20	73	7.2	453	1	CP51_WHEAT
21	73	7.2	797	1	SYM_ARATH
22	72.5	7.1	330	1	T2C9_CITFR
23	72.5	7.1	437	1	K1CM_MOUSE
24	72.5	7.1	440	1	MESJ_BUCAI
25	72.5	7.1	720	1	YD71_SCHPO
26	72	7.1	538	1	CK11_YEAST
27	72	7.1	928	1	RRPO_HUMAN
28	72	7.1	2210	1	RBPQ_LYCVA
29	71.5	7.0	539	1	YAH2_YEAST
30	71	7.0	161	1	FAB1_CAEEL
31	71	7.0	198	1	RS4_BACST
32	71	7.0	344	1	GCH2_PICGU
33	71	7.0	735	1	YDD7_SCHPO
34	71	7.0	1508	1	A10A_MOUSE
35	70.5	6.9	2555	1	FAFY_HUMAN
36	70.5	6.9	275	1	VG51_HSVSA
37	70.5	6.9	276	1	RK2_PICAB
38	70.5	6.9	357	1	RK2_PINTH
39	70.5	6.9	364	1	GGPP_CATRO
40	70.5	6.9	529	1	RRCF_AZOVI
41	70.5	6.9	846	1	SWA_DROPS
42	70	6.9	300	1	YDSB_SCHPO
43	70	6.9	700	1	Y43E_MYCPN
44	70	6.9	980	1	BIB_DROME
45	70	6.9	1508	1	POLG_LIV

RESULT 1
POOL_ECOLI
ID POOL_ECOLI STANDARD; PRT; 931 AA.
AC P31828; P31829; P76132; P78158;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable zinc protease pqqL (EC 3.4.99.-).
GN PQQ OR B1494.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RA Turlin E., Gasser F., Biville F.;
RT "Sequence and functional analysis of an Escherichia coli DNA fragment
able to complement pqqE and pqqF from Methylobacterium organophilum.";
RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R. Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RA MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
Sampei G., Seki Y., Sivasubraman S., Tagami H., Takeda J.,
Takenoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [4]
RP SIMILARITY: STRONG, TO H1N1INFLUENZA H1N1368.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M16.
CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
FRAMESHIFT IN POSITION 651 THAT PRODUCES TWO SEPARATE ORFS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

CC

```

CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HELICASE FAMILY. UVRD SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32811; AAC22966.1;
CC HSPG: P09980; IUAA.
CC TIGR: H11321;
CC InterPro: IPR004586; RecB.
CC InterPro: IPR000212; Uvrd-helicase.
CC Pfam: PF00580; Uvrd-helicase; 1.
CC TIGRFAMS: TIGR00609; recB; 1.
CC HydroLase: Nuclease; Exonuclease; Helicase; ATP-binding;
CC DNA repair; Complete proteome.
CC NP_BIND 22
CC FT ATP (POTENTIAL).
CC ST SEQUENCE 1211 AA; 139857 MW; 37B76430651FAD68 C64;
CC -----
CC Query Match 8.5%; Score 86.5; DB 1; Length 1211;
CC Best Local Similarity 20.9%; Pred. No. 2.3;
CC Matches 39; Conservative 33; Mismatches 80; Indels 35; Gaps 7;
CC -----
CC QY 11 LLTATSSLLTSN-IFCDELV-----MSNLHSKENYDKYSPRGYPKGER 55
CC DB 900 LEALPTLSINTKNDKDFASEFTGNIQDWRTSFTSEQAHRONY--FTE----SAGKK 953
CC QY 56 SLNFEELKDWGPKVNIKMTSPAVNKPMSFANPL-----RFGNVOEERSAGATA 106
CC DB 954 HAVFDDAKDYDSNAIEISTALLNENESNILDLPGRQVGTALHRHFENCYFSDLANTEE 1013
CC QY 107 NLPLRSGRNMESVLRVRRPNLPORFGRTTAKSVCRMLSDLCOGSMHSPCANDL-FYSMT 165
CC DB 1014 IDKLRFQSLQDQETTESLQNLQIQSHPLPNEIGIALADLA----NKDCEKEMPPYLAI 1069
CC QY 166 CQHQEIQ 172
CC DB 1070 REHFDE 1076
CC -----
CC RESULT 3
CC SNWA_DICDI
CC ID SNWA_DICDI STANDARD; PRT; 685 AA.
CC AC P54705;
CC DT 01-OCT-1996 (Rel. 34, Created)
CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
CC DT 01-NOV-1997 (Rel. 35, Last annotation update)
CC DE SNWA protein.
CC GN Dictyostelium discoideum (slime mold).
CC OS Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
CC OX NCBI TaxID=44689;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=97128797; PubMed=8973337;
CC RA Folk P., Puta F., Krpejsova L., Blahuskova A., Markos A.,
CC RA Rabino M., Dottin R.P.;
CC RT "The homolog of chromatin binding protein Bx42 identified in
CC RT Dictyostelium.";
CC RL Gene 181:229-231(1996).
CC -1- SIMILARITY: BELONGS TO THE SNW FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```


Query Match	7.9%;	Score 80;	DB 1;	Length 1086;
Best Local Similarity	27.9%;	Pred. No. 8.6;		
Matches	38;	Conservative	20;	Mismatches 36; Indels 42; Gaps 11;

RESULT 8			
PIV2_ADE05	PIV2_ADE05	STANDARD;	PRT; 449 AA
ID	PIV2_ADE05		
AC	P03271;		
DT	21-JUL-1986	(Rel. 01, Created)	
DT	21-JUL-1986	(Rel. 01, Last sequence update)	

RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
 RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
 RA Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
 RA Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,
 RA Wilcox L., Wohlman P., Vaudin M., Wilson R., Waterston R.,
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 [2]
 RN SEQUENCE OF 504-609 FROM N.A.
 RP STRAIN=5288c;
 RC MEDLINE=95116323; PubMed=7816623;
 RX Lygrou Z., Conesa C., Lesage P., Swanson R.N., Ruet A., Carlson M.,
 RA Sentenac A., Seraphin B.,
 RT "The yeast Bp1 gene encodes a transcription factor involved in the
 expression of a broad class of genes including snRNAs.";
 RL Nucleic Acids Res. 22:5332-5340(1994).
 CC - SIMILARITY: BELONGS TO THE UPF0034 (NIFR3/SMML) FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U19729; AAB82358.1; -;
 DR EMBL; 218944; CAA79376.1; -;
 DR SGD; S0004393; YLR401C.
 DR InterPro; IPR001269; UPF0034.
 DR Pfam; PF00642; zf-CCCH; 1.
 DR Pfam; PF01207; UPF0034; 1.
 DR PROSITE; PS01136; UPF0034; 1.
 KW Hypothetical protein.
 FT CONFLICT 602 602 W -> M (IN REF. 2).
 FT SEQUENCE 609 AA; 69817 MW; 948239C5B0F75909 CRC64;
 SQ
 Query Match 7.8%; Score 79; DB 1; Length 609;
 Best Local Similarity 21.7%; Pred. No. 5.1;
 Matches 35; Conservative 25; Mismatches 69; Indels 32; Gaps 4;
 QY 29 ELVMSNLHSHKENVKDYSEPRGYPKGRSLNFELKDWGPKNVKIKMTSPAVNKMPSFANL 88
 DB 261 ELSEHRMKQREVLKYLKDYRFAQEKPLDLVHKKTVSLTTV-----GNL 306
 QY 89 PLR-----FCRNVQERSAGATANPLRSGRNMEVSLVRVPNLPORFG-----R 133
 DB 307 PYRLMKKLGADVTYSEMALA---VPLIOGTNSEWALPKAHTSEFFGFGVQVACSKAWQA 363
 QY 134 TTTAKSVCRMSLDLCOGSMHSPCANDLFYSMTCHQHEIQNP 174
 DB 364 AKAAEALANSVSEITSEINLSGCPIDLLYRQSGSALLDNP 404
 RESULT 10
 PIV2_ADE07
 ID PIV2_ADE07 STANDARD; PRT; 448 AA.
 AC P03273; Q64833;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DE 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Maturational protein (Protein IVA2).
 GN PIVA2.
 OS Human adenovirus type 7.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OX NCBI_TaxID=10519;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83054637; PubMed=6292051;
 RA Engler J.A., van Bree M.P.;
 RT "The nucleotide sequence of the gene encoding protein Iva2 in human
 RT adenovirus type 7.";
 RL Gene 19:71-80(1982).

[2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Gomen;
 RA van Ormondt H.;
 RL Submitted (JUN-1985) to the EMBL/GenBank/DBJ databases.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X03000; CAA26766.1; -;
 DR PIR; A03843; Q4ADA7.
 DR InterPro; IPR003389; Adeno_IVa2.
 DR Pfam; PF02456; Adeno_IVa2; 1.
 KW ATP-binding.
 FT NP_BIND 171 178 ATP (POTENTIAL).
 FT CONFLICT 306 K -> R (IN REF. 2).
 SQ SEQUENCE 448 AA; 50678 MW; 01F17E29320848BC CRC64;
 Query Match 7.6%; Score 77; DB 1; Length 448;
 Best Local Similarity 23.4%; Pred. No. 5.5;
 Matches 39; Conservative 26; Mismatches 68; Indels 34; Gaps 8;
 QY 29 ELVMSNLHSHKENVKDYSEPRGYPKGRSLNFELKDWGPKNVKIKMTSPAVNKM 81
 DB 245 KWAYDDLTOEHNYD-VSDPRNV-----FARAAHGPPIAIIMDECMENLGCHKGVSKF 295
 QY 82 PHSF-ANLPLRFR-----NVQERSAGAT-ANPLRSGRNMEVSLVRVPNLP 128
 DB 296 FHAFPSKLHDKPKCTGYTVLVVHNNPRDLGNIANLKIQS--KMHIISPRMHPSQL 353
 QY 129 QRFGRITTT---AKSVCRMSLDLCOGSMHSPCANDLFYSMTCHQHEIQ 172
 DB 354 NRPVNTYTKGLPVAISLLKLDIVOHHALRCPYDQWVINYNTTPEALQ 400
 RESULT 11
 PIV2_ADE40
 ID PIV2_ADE40 STANDARD; PRT; 446 AA.
 AC P48752;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Maturational protein (Protein IVA2).
 GN PIVA2.
 OS Human adenovirus type 40.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OX NCBI_TaxID=28284;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=Dugan;
 RX MEDLINE=94087748; PubMed=8263936;
 RA Davison A.J., Telford E.A., Watson M.S., McBride K., Mautner V.;
 RT "The DNA sequence of adenovirus type 40.";
 RL J. Mol. Biol. 234:1308-1316(1993).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; LJ9443; AAC13952.1; -;
 DR InterPro; IPR003389; Adeno_IVa2.
 DR Pfam; PF02456; Adeno_IVa2; 1.
 KW ATP-binding.
 FT NP_BIND 170 177 ATP (POTENTIAL).

DB 379 PKGHIVATSPFANPLPHIYKNPDSYDPDRFCGREDKAAGAFYSISFGGGRH 432

RESULT 15

MYSB_CABEL

ID MYSB_CABEL STANDARD; PRT; 1966 AA.

AC P02566;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Myosin heavy chain B (MHC B).

GN UNC-54 OR MYO-4.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=83273600; PubMed=6576334;

RA Karn J., Brenner S., Barnett L.;

RT "protein structural domains in the Caenorhabditis elegans unc-54

RT myosin heavy chain gene are not separated by introns.;"

RL Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).

RN [2]

RP SEQUENCE OF 850-1966 FROM N.A.

RX MEDLINE=82272395; PubMed=7202124;

RA McLachlan A.D., Karn J.;

RT "periodic charge distributions in the myosin rod amino acid sequence

RT match cross-bridge spacings in muscle.;"

RL Nature 299:226-231(1982).

RN [3]

RP SEQUENCE OF 1876-1966 FROM N.A.

RX MEDLINE=83232892; PubMed=6571695;

RA Willis N., Gesteland R.F., Karn J., Barnett L., Bolten S.,

RA Waterston R.H.;

RT "The genes sup-7 X and sup-5 III of C. elegans suppress amber

RT nonsense mutations via altered transfer RNA.;"

RL Cell 33:575-583(1983).

CC -1- FUNCTION: MUSCLE CONTRACTION.

CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2

CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)

CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).

CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.

CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING

CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,

CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.

CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY

CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.

CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT

CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE

CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED

CC SUBFRAGMENT (S2).

CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN

CC C. ELEGANS.

CC -1- WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.

CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

CC EMBL: J01050; AAA28124.1; -

CC EMBL: V01494; CAA24738.1; -

CC PIR: A02992; MKW.

CC HSSP: P08799; IMNO.

CC InterPro: IPR004009; Myosin_N.

CC InterPro: IPR002928; Myosin_tail.

CC InterPro: IPR001609; myosin_head.

DR Pfam: PF00063; myosin_head; 1.

DR Pfam: PF01576; Myosin_tail; 1.

DR Pfam: PF02736; Myosin_N; 1.

DR PRINTS: PR00193; MYOSINHEAVY.

DR PRODOM: PD000355; myosin_head; 1.

DR SMART: SM00242; MYSC; 1.

KW Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;

KW ATP-binding; Methylation; Alkylation; Multigene family.

FT DOMAIN 1 850 MYOSIN HEAD-LIKE.

FT DOMAIN 851 1164 COILED COIL (POTENTIAL).

FT DOMAIN 1165 1176 ALPHA-HELICAL TAILPIECE (S2).

FT DOMAIN 1165 1966 HINGE.

FT NP_BIND 177 184 LIGHT MEROMYOSIN (LMW).

FT DOMAIN 665 687 ATP (BY SIMILARITY).

FT DOMAIN 769 783 ACTIN-BINDING.

FT MOD_RES 128 128 METHYLATION (TRI-) (POTENTIAL).

FT MOD_RES 705 705 ALKYLATION (SH-1).

FT MOD_RES 715 715 ALKYLATION (SH-2).

FT CONFLICT 1337 1337 E -> R (IN REF. 2).

FT CONFLICT 1880 1880 I -> L (IN REF. 2).

SQ SEQUENCE 1966 AA; 225125 MW; B66F0BB2FE27B67F CRC64;

Query Match 7.3%; Score 74.5; DB 1; Length 1966;

Best Local Similarity 19.2%; Pred. No. 63;

Matches 47; Conservative 46; Mismatches 101; Indels 51; Gaps 10;

QY 1 MEIISKPLILLTLATSSLL--TSNIFCADELVMNLHSHKENYDKYSEPRGYPKGRSLN 58

DB 523 IELIEKPLGIISMLDEECIVPKATDLTLASKLVQHLGKHPNEKPKPKG-KQGEAHEA 581

QY 59 FEELK-----DWGPKNVIKMSTPAVKMKPHSEAN---LPLRFGNRVQERSA----- 102

DB 582 MRHYAGTVRYNCLNWLEKNKNDPLNDTVVSAMKQSKGNDLVEIWDQYTTQEEAAKAKEG 641

QY 103 GATANPLRSGRNNEVSLVR-----VPNLQRFGRITTA---- 137

DB 642 GGGGKKKKKSGSEPTVMCLYRESLNLMTMLNKTHTPHFIRCIIPNEKKKQSGMDAALVLN 701

QY 138 KSVCRMLSD---LCQSGMHSPCANDLF---YSMTQCQHOEIQNPDKQO-SRRLLFKKIDDA 190

DB 702 QLTGCVLEIGIRICRKGFPNRTLHPDFVQRYAILAAKEAKSDDDKCKCAEAIMSKLVNDG 761

QY 191 ELKQE 195

DB 762 SLSEE 766

Search completed: March 26, 2003, 10:14:30

Job time : 30 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2003, 10:16:10 ; Search time 16 Seconds
(without alignments)
719.365 Million cell updates/sec

Title: US-09-831-758A-8
Perfect score: 1018
Sequence: 1 MEIISKLFILLATSSLL.....KQSRLLFKKIDDAELKQEK 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 237916 seqs, 58723674 residues
Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1014	99.6	196	US-09-995-793A-29	Sequence 29, Appl
2	82.5	8.1	4019	US-09-854-133-425	Sequence 425, App
3	82.5	8.1	4019	US-09-738-973-425	Sequence 425, App
4	81	8.0	1362	US-09-815-242-14009	Sequence 14009, A
5	75.5	7.4	869	US-10-016-283-33	Sequence 33, Appl
6	75.5	7.4	869	US-09-817-487A-2	Sequence 2, Appl
7	74.5	7.3	1170	US-10-135-687-2	Sequence 2, Appl
8	73	7.2	280	US-09-796-138-20	Sequence 20, Appl
9	73	7.2	280	US-09-909-903-20	Sequence 20, Appl
10	73	7.2	648	US-09-815-242-10504	Sequence 10504, A
11	72.5	7.1	348	US-09-840-787-10	Sequence 10, Appl
12	72	7.1	781	US-09-469-522-37	Sequence 37, Appl
13	72	7.1	797	US-09-469-522-45	Sequence 45, Appl
14	72	7.1	816	US-09-469-522-4	Sequence 4, Appl
15	72	7.1	832	US-09-469-522-35	Sequence 35, Appl
16	72	7.1	851	US-09-469-522-33	Sequence 33, Appl
17	72	7.1	851	US-09-469-522-39	Sequence 39, Appl
18	72	7.1	859	US-09-469-522-43	Sequence 43, Appl
19	72	7.1	869	US-09-469-522-47	Sequence 47, Appl

20	72	7.1	871	10	US-09-469-522-49	Sequence 49, Appl
21	72	7.1	874	10	US-09-469-522-31	Sequence 31, Appl
22	72	7.1	895	10	US-09-469-522-29	Sequence 29, Appl
23	72	7.1	897	10	US-09-469-522-41	Sequence 41, Appl
24	72	7.1	928	10	US-09-758-007-3	Sequence 3, Appl
25	72	7.1	928	10	US-09-860-211-8	Sequence 8, Appl
26	72	7.1	928	10	US-09-469-522-2	Sequence 2, Appl
27	72	7.1	928	10	US-09-469-522-51	Sequence 51, Appl
28	71	7.0	147	9	US-09-738-626-4146	Sequence 4146, Ap
29	71	7.0	377	10	US-09-934-778-2	Sequence 2, Appl
30	70.5	6.9	273	10	US-09-815-242-11993	Sequence 11993, A
31	70	6.9	1508	9	US-10-024-623-35	Sequence 35, Appl
32	70	6.9	1508	12	US-10-002-769-15	Sequence 15, Appl
33	69.5	6.8	747	9	US-09-978-295A-459	Sequence 459, App
34	69.5	6.8	747	9	US-09-938-418-9	Sequence 9, Appl
35	69.5	6.8	747	9	US-09-978-697-459	Sequence 459, App
36	69.5	6.8	747	9	US-09-978-132A-459	Sequence 459, App
37	69.5	6.8	747	9	US-09-939-832A-459	Sequence 459, App
38	69.5	6.8	747	9	US-09-978-189-459	Sequence 459, App
39	69.5	6.8	747	9	US-10-028-072-426	Sequence 426, App
40	69.5	6.8	747	9	US-10-121-049-426	Sequence 426, App
41	69.5	6.8	747	9	US-10-121-904-426	Sequence 426, App
42	69.5	6.8	747	9	US-10-140-470-426	Sequence 426, App
43	69.5	6.8	747	9	US-10-175-746-426	Sequence 426, App
44	69.5	6.8	747	9	US-10-176-918-426	Sequence 426, App
45	69.5	6.8	747	9	US-10-176-921-426	Sequence 426, App

ALIGNMENTS

RESULT 1
US-09-995-793A-29
; Sequence 29, Application US/09995793A
; Publication No. US20030054446A1
; GENERAL INFORMATION:
; APPLICANT: Weber, Bernard H.F.
; APPLICANT: Stoehr, Heidi
; TITLE OF INVENTION: No. US20030054446A1 retinal-specific human proteins C7orf
; FILE REFERENCE: 033488-001
; CURRENT APPLICATION NUMBER: US/09/995,793A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/253,751
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; TYPE: PRT
; LENGTH: 196
; ORGANISM: Homo sapiens
US-09-995-793A-29

Query Match 99.6%; Score 1014; DB 9; Length 196;
Best Local Similarity 99.5%; Pred. No. 7.8e-103;
Matches 195; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY	1	MEIISKLFILLATSSLLTSNIFC	ADLVMSNLHSENYDKYSEPRGYPKGRSLNFE	60
DB	1	MEIISKLFILLATSSLLTSNIFC	ADLVMSNLHSENYDKYSEPRGYPKGRSLNFE	60
QY	61	ELKDWGPKNYIKSTPAVKMPSFANL	PLRFRGNVOERSAGATANLPLRSGRNMEVSL	120
DB	61	ELKDWGPKNYIKSTPAVKMPSFANL	PLRFRGNVOERSAGATANLPLRSGRNMEVSL	120
QY	121	VRRVNPQRFGRITTTAKSVCRMLSD	LCQSMHSPCANDLFYSMTQHOETQNDQKOSR	180
DB	121	VRRVNPQRFGRITTTAKSVCRMLSD	LCQSMHSPCANDLFYSMTQHOETQNDQKOSR	180
QY	181	RLLFKKIDDAELKQEK	196	
DB	181	RLLFKKIDDAELKQEK	196	

```

RESULT 2
US-09-854-133-425
; Sequence 425, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854.133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 425
; LENGTH: 4019
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-854-133-425

Query Match      8.1%; Score 82.5; DB 9; Length 4019;
Best Local Similarity 20.6%; Pred. No. 21;
Matches 34; Conservative 27; Mismatches 61; Indels 43; Gaps 5;

QY 28 DELVMSNLHSEKNDYKSEP-----RGYPKGRSLNFEELKDWGPKNVKIM 73
DB 3733 EDLVLSDISPKGVWDKILEPVACVRRKKSEMLQLFPAYLKGE-----DLFGL 3778

QY 74 STPAVNMKPHSF-----ANLPLRFRGNVQERSAGATANLPLRSGRNMVEVSLVRRVPNL 127
DB 3779 TVSAVARIAESLPGVEACENYFRYGRNPLMELPLAVNPTGCARSEPKMSAHVKR-----3833

QY 128 PORFGRRTTAKSVCRMVSLDQSGMHSKPCANDLFYSMTQHQEIQ 172
DB 3834 PHTLNSSTSKS-----FQSTVTGELNAPYSKQFVHSSKSSQYRKMK 3874

RESULT 3
US-09-738-973-425
; Sequence 425, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 425
; LENGTH: 4019
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-738-973-425

Query Match      8.1%; Score 82.5; DB 10; Length 4019;
Best Local Similarity 20.6%; Pred. No. 21;
Matches 34; Conservative 27; Mismatches 61; Indels 43; Gaps 5;

```

```

QY 28 DELVMSNLHSEKNDYKSEP-----RGYPKGRSLNFEELKDWGPKNVKIM 73
DB 3733 EDLVLSDISPKGVWDKILEPVACVRRKKSEMLQLFPAYLKGE-----DLFGL 3778

QY 74 STPAVNMKPHSF-----ANLPLRFRGNVQERSAGATANLPLRSGRNMVEVSLVRRVPNL 127
DB 3779 TVSAVARIAESLPGVEACENYFRYGRNPLMELPLAVNPTGCARSEPKMSAHVKR-----3833

QY 128 PORFGRRTTAKSVCRMVSLDQSGMHSKPCANDLFYSMTQHQEIQ 172
DB 3834 PHTLNSSTSKS-----FQSTVTGELNAPYSKQFVHSSKSSQYRKMK 3874

RESULT 4
US-09-815-242-14009
; Sequence 14009, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14009
; LENGTH: 1362
; TYPE: PRT
; ORGANISM: Salmonella typhi
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1362)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-14009

Query Match      8.0%; Score 81; DB 10; Length 1362;
Best Local Similarity 25.2%; Pred. No. 6.8;
Matches 29; Conservative 24; Mismatches 42; Indels 20; Gaps 6;

QY 65 WGPKNVVKMSTPAVNMKPH-----SFANLPLRFRGNVQ-----ERSAGATANLPLRS 112
DB 1130 WAPGSDMVQHPVLEKLPYIVLVDEFADLMTVCKVEELIARLAQKARAAGIHLVLT 1189

QY 113 GR---NMEVSLVRRVPNLPRQFGRRTTAKSVCRMVSLDQSGMHSKPCANDLFYS 163
DB 1190 QRPVSDVITGLIK--ANIPTRIAFTVSSKIDSTILD--QGGAESLLGMDMLYS 1240

RESULT 5
US-10-016-283-33

```

```
; Sequence 33, Application US/10016283
; Patent No. US20020164702A1
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al., David M.
; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS AND LIGANDS
; FILE REFERENCE: REG195-B-PCT-US
; CURRENT APPLICATION NUMBER: US/10/016,283
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US/09/077,955A
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: PCT/US96/20696
; PRIOR FILING DATE: 1996-12-13
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33
; LENGTH: 869
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-016-283--33

Query Match          7.4%; Score 75.5; DB 9; Length 869;
Best Local Similarity 20.1%; Pred. No. 14;
Matches 44; Conservative 37; Mismatches 69; Indels 69; Gaps 14;

QY      1 MEIIS-KLFILLTATSLTSNIFCA-----LYCRRRKQWKKKRESAAVTLTLPSELLDLRLHP 550
       :||| :||| ||| |::| ::|
Db     498 ISINSFAFVLTTIT-----LYCRRRKQWKKKRESAAVTLTLPSELLDLRLHP 550

QY      38 KENTDKYS---EPR----GYPKGERSLNFEEELKDWP---KNVIK MSTPAVNKMPHS-FA 86
       |:|   |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|
Db     551 NPMYQRMPDLLNPKLLSLEYPRN----NIETVRDIGEGAGFGARAPGL--LPYEPT 604

QY      87 NLPLRGRNVQEERSAGATANPLRSGRNMVESLVRRV---PNLPQRFGRTRTTAKSVCRM 144
       :::: :::: ||| |::| ::|
Db     605 MVAVKG---LKREASADMQDF-----QREAALMAEFEDNPIVKLLGVCAVGKPMCLLF 655

QY      145 S DLCOG-----SMHSPCANLDIFYMTCOHQEI QNP 174
       :|    :|    |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|
Db     656 EYMAYGDLEFLNRMSPTHVCSS--LSHSDDL SNRAQVS SP 692


RESULT 6
US-09-817-487A-2
; Sequence 2, Application US/09817487A
; Patent No. US20020150876A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020150876Alartis AG
; TITLE OF INVENTION: Selectable Marker Genes
; FILE REFERENCE: 4-31193A
; CURRENT APPLICATION NUMBER: US/09/817,487A
; CURRENT FILING DATE: 2002-02-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 869
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-817-487A-2

Query Match          7.4%; Score 75.5; DB 10; Length 869;
Best Local Similarity 20.1%; Pred. No. 14;
Matches 44; Conservative 37; Mismatches 69; Indels 69; Gaps 14;

QY      1 MEIIS-KLFILLTATSLTSNIFCA-----DEIVMNSLNHS 37
       :||| :||| ||| |::| ::|
Db     498 ISINSFAFVLTTIT-----LYCRRRKQWKKKRESAAVTLTLPSELLDLRLHP 550

QY      38 KENYDKYS---EPR----GYPKGERSLNFEEELKDWP---KNVIK MSTPAVNKMPHS-FA 86
       |:|   |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|  |:|
Db     551 NPMYQRMPDLLNPKLLSLEYPRN----NIETVRDIGEGAGFGARAPGL--LPYEPT 604

QY      87 NLPLRGRNVQEERSAGATANPLRSGRNMVESLVRRV---PNLPQRFGRTRTTAKSVCRM 144
```


Thu Mar 27 11:08:44 2003

Db 378 SKLLNDNFHMSLLACALEVVNATYSRSTSONLDSGTDLSPFWILNVLNLKAFDFYKVIE 437
Qy 49 GYPKGRSLNFEELK-----DWGPKNVIKMSTPAYNKKPHSPANLPLR 91
Db 438 SFIRAEGLTREMILKHLERCEHRIMESLAWLSDSPFLDLIKOSKDRGPTDHLSESACPL- 496
Qy 92 FGRNVQERSACATANLPLRSGR---NMEVSLVRVPNLPORFORTTTAKSVCRMLSDLC 148
Db 497 -----NLPLQNNHTAADMYLSPVRS---PKKGGSTTRVNSTANAETO-- 535
Qy 149 QGSMHSPCANDLFYSMTCOHOEIQNPOKQSRRLLEKKI 187
Db 536 -----ATSAF-----QOKPLKSTLSLSEYKKV 558

Search completed: March 26, 2003, 10:23:51
Job time : 20 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2003, 10:08:29 ; Search time 35 Seconds
(without alignments)
746.203 Million cell updates/sec

Title: US-09-831-758A-8
Perfect score: 1018
Sequence: 1 MEISSLKILLTATSSLL.....KQSRLLFKKIDDAELKQEK 196

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1018	100.0	196	21	Novel human G-prot
2	1018	100.0	196	22	Amino acid sequenc
3	1018	100.0	196	23	Human retina speci
4	1014	99.6	196	23	Human retina speci
5	940	92.3	180	21	Novel human G-prot
6	940	92.3	180	22	Amino acid sequenc
7	674	66.2	196	21	Novel bovine G-pro.
8	674	66.2	196	22	Amino acid sequenc
9	552.5	54.3	203	21	Novel rat G-protei
10	552.5	54.3	203	22	Amino acid sequenc

11	546.5	53.7	203	21	AAV93143	Novel rat G-protei
12	546.5	53.7	203	22	AAG67750	Amino acid sequenc
13	517	50.8	188	21	AAV93144	Novel mouse G-prot
14	517	50.8	188	22	AAG67751	Amino acid sequenc
15	98	9.6	354	21	AAV93145	Arabidopsis thalia
16	98	9.6	354	22	AAV93146	Arabidopsis thalia
17	98	9.6	426	23	AAV93147	Herbicidally activ
18	98	9.6	426	23	AAV93148	Herbicidally activ
19	88	8.6	1429	22	ABV58779	Arabidopsis thalia
20	85	8.3	1436	22	ABV58780	Drosophila melanog
21	83.5	8.2	1436	22	ABV58781	Drosophila melanog
22	83.5	8.2	1436	22	ABV58782	Human protein sequ
23	82.5	8.1	408	21	ABP27526	Streptococcus poly
24	82.5	8.1	4019	22	AAE13839	Human lung tumour-
25	81.5	8.0	1435	22	ABV11808	Human BCI9 homolog
26	81	8.0	1362	22	AAU38416	Salmonella typhi c
27	78	7.7	753	22	ABV60959	Drosophila melanog
28	77	7.6	775	23	AAU93161	Arabidopsis thalia
29	76.5	7.5	530	17	AAV92715	Mouse muscle-local
30	76.5	7.5	860	17	AAV92716	Mouse muscle-local
31	76.5	7.5	868	17	AAV92717	Mouse muscle-local
32	76	7.5	296	21	AAV92718	Arabidopsis thalia
33	76	7.5	313	21	AAV92719	Arabidopsis thalia
34	75.5	7.4	371	21	AAV92720	Arabidopsis thalia
35	75.5	7.4	552	22	ABV1902	Drosophila melanog
36	75.5	7.4	869	18	AAV26611	Human muscle-speci
37	75.5	7.4	869	18	AAV26612	Human Dmk receptor
38	75.5	7.4	869	22	AAV77856	Protein of muscle
39	75.5	7.4	869	22	AAV68421	Amino acid sequenc
40	75	7.4	296	21	AAV52208	Arabidopsis thalia
41	75	7.4	313	21	AAV52209	Arabidopsis thalia
42	75	7.4	517	23	ABV53402	Lactococcus lactis
43	75	7.4	725	23	AAE14506	Human mitotic cent
44	75	7.4	928	16	AAV71680	Retinoblastoma ppl
45	75	7.4	928	16	AAV71681	Recombinant ppl10R

ALIGNMENTS

RESULT 1
AAV93141
ID AAV93141 standard; Protein; 196 AA.
AC AAV93141;
DT 06-DEC-2000 (first entry)
DE Novel human G-protein coupled receptor #2.
KW G-protein coupled receptor; human; bovine; nervous system disorder;
KW rat; mouse; somatostatin excretion.
OS Homo sapiens.
PN WO200029441-A1.
PD 25-MAY-2000.
PF 11-NOV-1999; 99MO-JP06283.
PR 13-NOV-1998; 98JP-0323759.
PR 08-MAR-1999; 99JP-0060030.
PR 14-APR-1999; 99JP-0106812.
PR 14-JUN-1999; 99JP-0166672.
PR 04-AUG-1999; 99JP-0221640.
PR 14-SEP-1999; 99JP-0259818.
(TAKE) TAKEDA CHEM IND LTD.
XX Watanabe T, Kikuchi K, Terao Y, Shintani Y, Hinuma S, Fukusumi S;
XX Fujii R, Hosoya M, Kitada C;
XX

DR WPI; 2000-387747/33.
 DR N-PSDB; AAA70501.
 XX G protein coupled-receptor protein and antibodies to it for treatment
 PT and diagnosis of nerve diseases
 PT
 PS Claim 2; Fig 3; 184pp; Japanese.
 XX
 CC The invention relates to the isolation of novel G-protein coupled
 CC receptor (GPCR) genes and their encoded proteins. This sequence
 CC represents the protein sequence of a human GPCR. The DNAs and proteins
 CC are used for the treatment, prevention and diagnosis of disorders of
 CC the nervous system. The proteins and its fragments are also promoters
 CC of somatostatin excretion.
 XX
 SQ Sequence 196 AA;
 Query Match 100.0%; Score 1018; DB 21; Length 196;
 Best Local Similarity 100.0%; Pred. No. 5e-106;
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MEIISKLFILLTATSSLLTSNIFCADELVMSNLHSENKYDSEPRGYPKGRSLNFE 60
 Db 1 MEIISKLFILLTATSSLLTSNIFCADELVMSNLHSENKYDSEPRGYPKGRSLNFE 60
 Qy 61 ELKDWGPNVKIMSTPAVNKMPHSFANPLRFRGNVQERSAGATANPLRSGRNMEVSL 120
 Db 61 ELKDWGPNVKIMSTPAVNKMPHSFANPLRFRGNVQERSAGATANPLRSGRNMEVSL 120
 Qy 121 VRRVNLPRQFGRTTAKSVCRMLSDLCQSGSMHSPCANDLFYSMTCHQHOEIONPDQKQR 180
 Db 121 VRRVNLPRQFGRTTAKSVCRMLSDLCQSGSMHSPCANDLFYSMTCHQHOEIONPDQKQR 180
 Qy 181 RLLFKKIDDAELKQEK 196
 Db 181 RLLFKKIDDAELKQEK 196
 RESULT 2
 AAG67748
 ID AAG67748 standard; Protein; 196 AA.
 AC AAG67748;
 XX
 XX 10-DEC-2001 (first entry)
 DT
 DE Amino acid sequence of a human prolactin secretion regulating agent.
 XX Human; prolactin secretion; hypooovarianism; sperm development;
 KW osteoporosis; lactation disorder; hypothyroidism; kidney failure;
 KW menopause; hyperprolactinemia; pituitary tumour; dienecephalon tumour;
 KW menstrual disorder; stress; autoimmune disease; prolactinoma; sterility;
 KW impotence; amenorrhea; lactorexia; hyperpituitarism; Sheehan's syndrome
 KW Chari-Frommel syndrome; Argonz-del Castillo syndrome; lymphoma;
 KW Forbes-Albright syndrome; spermatogenesis disorder.
 XX
 OS Homo sapiens.
 XX
 XX WO200166134-A1.
 PN
 XX 13-SEP-2001.
 PD
 XX 06-MAR-2001; 2001WO-JP01716.
 PF
 XX 06-MAR-2000; 2000JP-0065752.
 PR
 XX 07-DEC-2000; 2000JP-0378001.
 XX
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA Matsumoto Y, Watanabe T, Hinuma S, Habata Y, Yoshida H;
 PI WPI: 2001-596812/67.
 DR N-PSDB; AAH78481.
 DR

XX New polypeptide, useful as a vasotropic, tranquilizing,
 PT immunosuppressive and gynecological agent comprises the prolactin
 PT secretion regulator
 PT
 XX Claim 2; Fig 3; 180pp; Japanese.
 PS
 XX The present sequence represents a human polypeptide which is a prolactin
 CC secretion regulating agent. The prolactin secretion regulating agent
 CC polypeptide and polynucleotide are used for the treatment and prevention
 CC of hypooovarianism, sperm under development, osteoporosis, menopause,
 CC lactation disorders, hypothyroidism, kidney failure, hyperprolactinemia,
 CC pituitary tumour, dienecephalon tumour, menstrual disorders, stress,
 CC autoimmune disease, prolactinoma, sterility, impotence, amenorrhea,
 CC lactorexia, hyperpituitarism, Chari-Frommel syndrome, Argonz-del
 CC Castillo syndrome, Forbes-Albright syndrome, lymphoma, Sheehan's syndrome
 CC and spermatogenesis disorders.
 XX
 SQ Sequence 196 AA;
 Query Match 100.0%; Score 1018; DB 22; Length 196;
 Best Local Similarity 100.0%; Pred. No. 5e-106;
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MEIISKLFILLTATSSLLTSNIFCADELVMSNLHSENKYDSEPRGYPKGRSLNFE 60
 Db 1 MEIISKLFILLTATSSLLTSNIFCADELVMSNLHSENKYDSEPRGYPKGRSLNFE 60
 Qy 61 ELKDWGPNVKIMSTPAVNKMPHSFANPLRFRGNVQERSAGATANPLRSGRNMEVSL 120
 Db 61 ELKDWGPNVKIMSTPAVNKMPHSFANPLRFRGNVQERSAGATANPLRSGRNMEVSL 120
 Qy 121 VRRVNLPRQFGRTTAKSVCRMLSDLCQSGSMHSPCANDLFYSMTCHQHOEIONPDQKQR 180
 Db 121 VRRVNLPRQFGRTTAKSVCRMLSDLCQSGSMHSPCANDLFYSMTCHQHOEIONPDQKQR 180
 Qy 181 RLLFKKIDDAELKQEK 196
 Db 181 RLLFKKIDDAELKQEK 196
 RESULT 3
 AAU99162
 ID AAU99162 standard; Protein; 196 AA.
 XX
 AC AAU99162;
 XX
 XX 24-SEP-2002 (first entry)
 DT
 DE Human retina specific protein C7orf9 variant protein.
 XX Human; MPP4; C7orf9; C12orf7; F379; retina specific gene;
 KW AMD; age-related macular degeneration; blindness; gene therapy;
 KW Ophthalmological; transgenic.
 KW
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 32
 FT /note= "Wild-type Ile substituted by Met"
 ET
 XX WO200244366-A2.
 PN
 XX 06-JUN-2002.
 PD
 XX 29-NOV-2001; 2001WO-EP13940.
 PF
 XX 29-NOV-2000; 2000US-253751P.
 PR
 XX (MULT-) MULTIGENE BIOTECH GMBH.
 PA Stoehr HB, Weber BHF;
 PI
 XX

DR WPI: 2002-508512/54.
 XX Novel nucleic acids encoding retina-specific human protein C7orf9,
 PT C12orf7, MPP4 or F379, useful for diagnosing age-related macular
 PT degeneration or predisposition for macular degeneration, and in gene
 PT therapy techniques -
 XX Example 2; Page : 120pp; English.
 XX The invention relates to isolated nucleic acid encoding retina-specific
 CC human protein C7orf9, C12orf7, MPP4 or F379 or a fragment, derivative or
 CC allelic variation of the above mentioned nucleic acid sequences.
 CC Also included are a recombinant vector containing the nucleic acid, a
 CC recombinant host cell which contains the vector and expresses the
 CC protein, an inhibitor characterised in that it can suppress the
 CC activity of the protein, treating macular degeneration or a
 CC predisposition for macular degeneration, comprising administering to a
 CC mammalian subject a reagent which decreases, inhibits or increases
 CC expression of C7orf9, C12orf7, MPP4 and/or F379 or which leads to the
 CC expression of a biologically active C7orf9, C12orf7, MPP4 and/or F379
 CC protein and a transgenic non-human animal comprising at least one of the
 CC nucleic acids (active or inactivated). The nucleic acid or protein is
 CC useful for diagnosing macular degeneration, preferably age-related
 CC macular degeneration (AMD) or a predisposition for macular degeneration.
 CC The reagent used in the diagnosis is a C7orf9-, C12orf7-, MPP4- or F379
 CC -specific nucleic acid probe, or anti-C7orf9, anti-C12orf7, anti-MPP4 or
 CC anti-F379-antibody. The reagent is detectably labeled, with a compound
 CC such as a radioisotope, a bioluminescent compound, a chemiluminescent
 CC compound, a fluorescent compound, a metal chelate or an enzyme. Fragments
 CC of the nucleic acid are useful as probes or primers in a diagnostic
 CC assay, and for identifying further factors involved in development and
 CC progression of macular degeneration. The proteins encoded by the nucleic
 CC acid are useful to identify further unrelated proteins which are
 CC associated with macular degeneration and for use in screening methods
 CC based on protein/protein interactions. The nucleic acid is also useful as
 CC reagents for detecting differences between normal and aberrant expression
 CC of the protein. The nucleic acid is also useful in gene therapy
 CC techniques, and can be used for gene targeting and/or gene replacement
 CC for restoring a mutant gene or for creating a mutant gene via homologous
 CC recombination. The protein can be used to identify other proteins
 CC involved in development and progression of macular degeneration.
 CC The present sequence represents a variant of the retina specific protein
 CC C7orf9.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the sequence appearing as ABG66919 and
 CC the information in example 2.

XX Sequence 196 AA;
 Query Match 100.0%; Score 1018; DB 23; Length 196;
 Best Local Similarity 100.0%; Pred. No. 5e-106;
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEITSSKFLILLTATSSLLTSNIFCADELVMNLSKENVKDYSEPRGYPKGRSLNFE 60
 Db 1 MEITSSKFLILLTATSSLLTSNIFCADELVMNLSKENVKDYSEPRGYPKGRSLNFE 60
 QY 61 ELKDWGPKNVTKMSTPAVNMKPHSFANLPLRFGNRVQERSAGATANLPLRSGRNMEVSL 120
 Db 61 ELKDWGPKNVTKMSTPAVNMKPHSFANLPLRFGNRVQERSAGATANLPLRSGRNMEVSL 120
 QY 121 VRRVPLNLPQRFGRRTTAKSVCRMLSDLCQSMHSPCANLFLYSMTQCHQEQIONPDQKOSR 180
 Db 121 VRRVPLNLPQRFGRRTTAKSVCRMLSDLCQSMHSPCANLFLYSMTQCHQEQIONPDQKOSR 180
 QY 181 RLLEFKKIDDAELKQEK 196
 Db 181 RLLEFKKIDDAELKQEK 196

RESULT 4
 ABG66919
 ID ABG66919 standard; Protein: 196 AA.

XX ABG66919;
 XX 24-SEP-2002 (first entry)
 XX Human retina specific protein encoded by cDNA C7orf9.
 XX Human; MPP4; C7orf9; C12orf7; F379; retina specific gene;
 KW AMD; age-related macular degeneration; blindness; gene therapy;
 KW Ophthalmological; transgenic.
 XX Homo sapiens.
 XX WO200244366-A2.
 XX 06-JUN-2002.
 XX 29-NOV-2001; 2001WO-EP13940.
 XX 29-NOV-2000; 2000US-253751P.
 XX (MULT-) MULTIGENE BIOTECH GMBH.
 XX Stoehr HB, Weber BHF;
 XX WPI: 2002-508512/54.
 XX N-PSDB; ABK95399.
 XX Novel nucleic acids encoding retina-specific human protein C7orf9,
 PT C12orf7, MPP4 or F379, useful for diagnosing age-related macular
 PT degeneration or predisposition for macular degeneration, and in gene
 PT therapy techniques -
 XX Claim 1; Fig 9; 120pp; English.
 XX The invention relates to isolated nucleic acid encoding retina-specific
 CC human protein C7orf9, C12orf7, MPP4 or F379 or a fragment, derivative or
 CC allelic variation of the above mentioned nucleic acid sequences.
 CC Also included are a recombinant vector containing the nucleic acid, a
 CC recombinant host cell which contains the vector and expresses the
 CC protein, an inhibitor characterised in that it can suppress the
 CC activity of the protein, treating macular degeneration or a
 CC predisposition for macular degeneration, comprising administering to a
 CC mammalian subject a reagent which decreases, inhibits or increases
 CC expression of C7orf9, C12orf7, MPP4 and/or F379 or which leads to the
 CC expression of a biologically active C7orf9, C12orf7, MPP4 and/or F379
 CC protein and a transgenic non-human animal comprising at least one of the
 CC nucleic acids (active or inactivated). The nucleic acid or protein is
 CC useful for diagnosing macular degeneration, preferably age-related
 CC macular degeneration (AMD) or a predisposition for macular degeneration.
 CC The reagent used in the diagnosis is a C7orf9-, C12orf7-, MPP4- or F379
 CC -specific nucleic acid probe, or anti-C7orf9, anti-C12orf7, anti-MPP4 or
 CC anti-F379-antibody. The reagent is detectably labeled, with a compound
 CC such as a radioisotope, a bioluminescent compound, a chemiluminescent
 CC compound, a fluorescent compound, a metal chelate or an enzyme. Fragments
 CC of the nucleic acid are useful as probes or primers in a diagnostic
 CC assay, and for identifying further factors involved in development and
 CC progression of macular degeneration. The proteins encoded by the nucleic
 CC acid are useful to identify further unrelated proteins which are
 CC associated with macular degeneration and for use in screening methods
 CC based on protein/protein interactions. The nucleic acid is also useful as
 CC reagents for detecting differences between normal and aberrant expression
 CC of the protein. The nucleic acid is also useful in gene therapy
 CC techniques, and can be used for gene targeting and/or gene replacement
 CC for restoring a mutant gene or for creating a mutant gene via homologous
 CC recombination. The protein can be used to identify other proteins
 CC involved in development and progression of macular degeneration.
 CC The present sequence represents a retina specific protein of
 CC the invention.
 XX Sequence 196 AA;

Query Match 99.6%; Score 1014; DB 23; Length 196;

Best Local Similarity 99.5%; Pred. No. 1.4e-105;
Matches 195; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIISKFLITLATSSLLTSNIFCADELVMSNLHSEKYNKYSEPRGYPKGRSLNPE 60
DB 1 MEIISKFLITLATSSLLTSNIFCADELVMSNLHSEKYNKYSEPRGYPKGRSLNPE 60
QY 61 ELKDMGPKNVIKMSTPAVNKMPHSFANLPLRFGNNVQERSAGATANLPLRSGRNNEVSL 120
DB 61 ELKDMGPKNVIKMSTPAVNKMPHSFANLPLRFGNNVQERSAGATANLPLRSGRNNEVSL 120
QY 121 VRRVPLNLPORFGRTTTAKSVCRMLSDLCOGSMHSPCANDLFYSMTCHQHOEIQNPDQKQR 180
DB 121 VRRVPLNLPORFGRTTTAKSVCRMLSDLCOGSMHSPCANDLFYSMTCHQHOEIQNPDQKQR 180
QY 181 RLLFKKIDDAELKQEK 196
DB 181 RLLFKKIDDAELKQEK 196

RESULT 5
AAY93140
ID AAY93140 standard; Protein; 180 AA.
XX
AC AAY93140;
XX
DT 06-DEC-2000 (first entry)
XX
DE Novel human G-protein coupled receptor #1.
XX
KW G-protein coupled receptor; human; bovine; nervous system disorder;
KW rat; mouse; somatostatin excretion.
XX
OS Homo sapiens.
XX
PN WO200029441-A1.
XX
PD 25-MAY-2000.
XX
PF 11-NOV-1999; 99WO-JP06283.
XX
PR 13-NOV-1998; 98JP-0323759.
PR 08-MAR-1999; 99JP-0060030.
PR 14-APR-1999; 99JP-0106812.
PR 14-JUN-1999; 99JP-0166672.
PR 04-AUG-1999; 99JP-0221640.
PR 14-SEP-1999; 99JP-0259818.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Watanabe T, Kikuchi K, Terao Y, Shintani Y, Hinuma S, Fukusumi S;
PI Fujii R, Hosoya M, Kitada C;
XX
DR WPI: 2000-387747/33.
DR N-PSDB; AAY70500.
XX
G protein coupled receptor protein and antibodies to it for treatment
and diagnosis of nerve diseases
XX
PS Claim 1; Fig 1; 184pp; Japanese.
XX
CC The invention relates to the isolation of novel G-protein coupled
CC receptor (GPCR) genes and their encoded proteins. This sequence
CC represents the protein sequence of a human GPCR. The DNAs and proteins
CC are used for the treatment, prevention and diagnosis of disorders of
CC the nervous system. The proteins and its fragments are also promoters
CC of somatostatin excretion.
XX
SQ Sequence 180 AA;

Query Match 92.3%; Score 940; DB 21; Length 180;
Best Local Similarity 100.0%; Pred. No. 2.6e-97;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIISKFLITLATSSLLTSNIFCADELVMSNLHSEKYNKYSEPRGYPKGRSLNPE 60
DB 1 MEIISKFLITLATSSLLTSNIFCADELVMSNLHSEKYNKYSEPRGYPKGRSLNPE 60
QY 61 ELKDMGPKNVIKMSTPAVNKMPHSFANLPLRFGNNVQERSAGATANLPLRSGRNNEVSL 120
DB 61 ELKDMGPKNVIKMSTPAVNKMPHSFANLPLRFGNNVQERSAGATANLPLRSGRNNEVSL 120
QY 121 VRRVPLNLPORFGRTTTAKSVCRMLSDLCOGSMHSPCANDLFYSMTCHQHOEIQNPDQKQR 180
DB 121 VRRVPLNLPORFGRTTTAKSVCRMLSDLCOGSMHSPCANDLFYSMTCHQHOEIQNPDQKQR 180

RESULT 6
AAG67747
ID AAG67747 standard; Protein; 180 AA.
XX
AC AAG67747;
XX
DT 10-DEC-2001 (first entry)
XX
DE Amino acid sequence of a human prolactin secretion regulating agent.
XX
KW Human; prolactin secretion; hypovarianism; sperm development;
KW osteoporosis; lactation disorder; hypothyroidism; kidney failure;
KW hyperprolactinemia; pituitary tumour; dienecephalon tumour; menopause;
KW menstrual disorder; stress; autoimmune disease; prolactinoma; sterility;
KW impotence; amenorrhea; lactorrhea; hyperpituitarism; Sheehan's syndrome
KW Chaiiri-Frommel syndrome; Argonz-del Castillo syndrome; lymphoma;
KW Forbes-Albright syndrome; spermatogenesis disorder.
XX
OS Homo sapiens.
XX
PN WO200166134-A1.
XX
PD 13-SEP-2001.
XX
PF 06-MAR-2001; 2001WO-JP01716.
XX
PR 06-MAR-2000; 2000JP-0065752.
PR 07-DEC-2000; 2000JP-0378001.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Matsumoto Y, Watanabe T, Hinuma S, Habata Y, Yoshida H;
PI WPI: 2001-596812/67.
DR N-PSDB; AAH78475.
XX
DR New polypeptide, useful as a vasotropic, tranquillizing,
PT immunosuppressive and gynecological agent comprises the prolactin
PT secretion regulator
XX
PS Claim 1; Fig 1; 180pp; Japanese.
XX
CC The present sequence represents a human polypeptide which is a prolactin
CC secretion regulating agent. The prolactin secretion regulating agent
CC polypeptide and polynucleotide are used for the treatment and prevention
CC of hypovarianism, sperm under development, osteoporosis, menopause,
CC lactation disorders, hypothyroidism, kidney failure, hyperprolactinemia,
CC pituitary tumour, dienecephalon tumour, menstrual disorders, stress,
CC autoimmune disease, prolactinoma, sterility, impotence, amenorrhea,
CC lactorrhea, hyperpituitarism, Chaiiri-Frommel syndrome, Argonz-del
CC Castillo syndrome, Forbes-Albright syndrome, lymphoma, Sheehan's syndrome
CC and spermatogenesis disorders.
XX
SQ Sequence 180 AA;

Query Match 92.3%; Score 940; DB 22; Length 180;
Best Local Similarity 100.0%; Pred. No. 2.6e-97;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 NPEELKMGKPNVKNMSTPAVNMKPHSFANPLRFRGNVQERSAGATANPLRSGRNM 117
 DB 60 TEEVKDAPK--IKMKPVNKPMPSSANPLRFRGNMEERSTRAMHPLRLGKNRE 117
 QY 118 VSLRRVNLPRFORGRTTAKSVCRMLSDLCOGSMHSPCANDLFYSMTQCHOETONPDQK 177
 DB 118 DLSLRVNLPRFORGRTTAKSVCRMLSDLCOGSMHSPCANDLFYSMTQCHOETONPDQK 177
 QY 178 QSRLLFKKIDDAELKQEK 196
 DB 178 NLRRGFKIDDAELKQEK 196

RESULT 9
 ID AAY93145 standard; Protein; 203 AA.
 AC AAY93145;
 XX
 DT 06-DEC-2000 (first entry)
 DE Novel rat G-protein coupled receptor #2.
 XX
 KW G-protein coupled receptor; human; bovine; nervous system disorder;
 KW rat; mouse; somatostatin excretion.
 XX
 OS Rattus sp.
 XX
 PN WO200029441-A1.
 XX
 PD 25-MAY-2000.
 XX
 PF 11-NOV-1999; 99WO-JP06283.
 XX
 PR 13-NOV-1998; 98JP-0323759.
 PR 08-MAR-1999; 98JP-0060030.
 PR 14-APR-1999; 98JP-0108812.
 PR 14-JUN-1999; 99JP-0166672.
 PR 04-AUG-1999; 99JP-0221640.
 PR 14-SEP-1999; 99JP-0259818.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Watanabe T, Kikuchi K, Terao Y, Shintani Y, Hinuma S, Fukusumi S;
 PI Fujii R, Hosoya M, Kitada C;
 XX
 DR WPI; 2000-387747/33.
 DR N-PSDB; AAA70505.
 XX
 PT G protein coupled receptor protein and antibodies to it for treatment
 PT and diagnosis of nerve diseases
 XX
 PS Claim 2; Page 176-177; 184pp; Japanese.

CC The invention relates to the isolation of novel G-protein coupled
 CC receptor (GPCR) genes and their encoded proteins. This sequence
 CC represents the protein sequence of a rat GPCR. The DNAs and proteins
 CC are used for the treatment, prevention and diagnosis of disorders of
 CC the nervous system. The proteins and its fragments are also promoters
 CC of somatostatin excretion.
 XX
 SQ Sequence 203 AA;
 Query Match 54.3%; Score 552.5; DB 21; Length 203;
 Best Local Similarity 58.1%; Pred. No. 1.2e-53;
 Matches 115; Conservative 26; Mismatches 46; Indels 11; Gaps 3;
 QY 1 MEIISKLFIILLATSSLLTSNIFCADELVMNLHKNYDKYSEPRGPKG--ERSLN 58
 DB 1 MEIISKRFILLATSSLLTSNIFCADELVMNLHKNYDKYSEPRGPKG--ERSLN 58
 QY 59 FEELKMGKPNVKNMSTPAVNMKPHSFANPLRFRGNVQERSAGATANPLRSGRNM 118

DB 61 FOELKMGKADIKMSPAPANKVPHSAANPLRFRGNIEDRRSPARA-----NMEA 112
 QY 119 SLVRRVNLPRFORGRTTAKSVCRMLSDLCOGSMHSPCANDLFYSMTQCHOETONPDQK 178
 DB 113 GTMHHFSLPRFORG--TTARRITKTLAGLPKSLHSLASSELLYAMTRQHQEIQSGOEQ 171
 QY 179 SRRLLFKKIDDAELKQEK 196
 DB 172 PKRVFTETDDAERKQEK 189

RESULT 10
 ID AAG67756 standard; Protein; 203 AA.
 AC AAG67756;
 XX
 DT 10-DEC-2001 (first entry)
 DE Amino acid sequence of a rat prolactin secretion regulating agent.
 XX
 KW Rat; prolactin secretion; hypovarianism; sperm development;
 KW osteoporosis; lactation disorder; hypothyroidism; kidney failure;
 KW menopause; hyperprolactinemia; pituitary tumour; dienecephalon tumour;
 KW menstrual disorder; stress; autoimmune disease; prolactinoma; sterility;
 KW impotence; amenorrhea; lactorrhea; hyperpituitarism; Sheehan's syndrome
 KW Chairi-Frommel syndrome; Argonz-del Castillo syndrome; lymphoma;
 KW Forbes-Albright syndrome; spermatogenesis disorder.
 XX
 OS Rattus sp.
 XX
 PN WO200166134-A1.
 XX
 PD 13-SEP-2001.
 XX
 PF 06-MAR-2001; 2001WO-JP01716.
 XX
 PR 06-MAR-2000; 2000JP-0065752.
 PR 07-DEC-2000; 2000JP-0378001.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Matsumoto Y, Watanabe T, Hinuma S, Habata Y, Yoshida H;
 XX
 DR WPI; 2001-596812/67.
 DR N-PSDB; AAH78515.
 XX
 PT New polypeptide, useful as a vasotropic, tranquilizing,
 PT immunosuppressive and gynecological agent comprises the prolactin
 PT secretion regulator
 XX
 PS Claim 2; Page 166-167; 180pp; Japanese.

CC The present sequence represents a rat polypeptide which is a prolactin
 CC secretion regulating agent. The prolactin secretion regulating agent
 CC polypeptide and polynucleotide are used for the treatment and prevention
 CC of hypovarianism, sperm under development, osteoporosis, menopause,
 CC lactation disorders, hypothyroidism, kidney failure, hyperprolactinemia,
 CC pituitary tumour, dienecephalon tumour, menstrual disorders, stress,
 CC autoimmune disease, prolactinoma, sterility, impotence, amenorrhea,
 CC lactorrhea, hyperpituitarism, Chairi-Frommel syndrome, Argonz-del
 CC Castillo syndrome, Forbes-Albright syndrome, lymphoma, Sheehan's syndrome
 CC and spermatogenesis disorders.
 XX
 SQ Sequence 203 AA;
 Query Match 54.3%; Score 552.5; DB 22; Length 203;
 Best Local Similarity 58.1%; Pred. No. 1.2e-53;
 Matches 115; Conservative 26; Mismatches 46; Indels 11; Gaps 3;
 QY 1 MEIISKLFIILLATSSLLTSNIFCADELVMNLHKNYDKYSEPRGPKG--ERSLN 58
 DB 1 MEIISKRFILLATSSLLTSNIFCADELVMNLHKNYDKYSEPRGPKG--ERSLN 58

Db 1 MEIISRRFILLTATSFLLSNTLCDELMPHFHSKSGYKYYQLRGIPKGVKERSVT 60
 Qy 59 FEELKDWGPKNVKIMSTPAVNKMPHSFANLPLRFGNNVOERSAGATANLPLRSGRNNEV 118
 Db 61 FOELKDWGAKKIDIKMSPAPANKVPHSAANLPLRFGNIEDRRSPARA-----NNEA 112
 Qy 119 SLVRVPLNLPORFGRTTAKSVCRMLSDLCOGSMHSPCANDLFYSMTCHQHOEIQNPQOK 178
 Db 113 GTSMHFFSLPQRFGR-TTARRITKTLAGLPQKSLHSLASSESLYAMTROHOEIQSPQEQ 171
 Qy 179 SRRLFKKIDDAELKQEK 196
 Db 172 PKRVFTETDDAERKQEK 189

RESULT 13
 AAY93144
 ID AAY93144 standard; Protein; 188 AA.
 AC AAY93144;
 XX
 DT 06-DEC-2000 (first entry)
 DE
 DE Novel mouse G-protein coupled receptor #1.
 KW G-protein coupled receptor; human; bovine; nervous system disorder;
 KW rat; mouse; somatostatin excretion.
 OS Mus sp.
 XX WO200029441-A1.
 XX 25-MAY-2000.
 XX 11-NOV-1999; 99WO-JP06283.
 XX 13-NOV-1998; 98JP-0323759.
 PR 08-MAR-1999; 99JP-0060030.
 PR 14-APR-1999; 99JP-0106812.
 PR 14-JUN-1999; 99JP-0166672.
 PR 04-AUG-1999; 99JP-0221640.
 PR 14-SEP-1999; 99JP-0259818.
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX Watanabe T, Kikuchi K, Terao Y, Shintani Y, Hinuma S, Fukusumi S;
 PI Fujii R, Hosoya M, Kitada C;
 XX WPI; 2000-387747/33.
 DR N-PSDB; AAA70504.
 XX G protein coupled receptor protein and antibodies to it for treatment
 PT and diagnosis of nerve diseases
 XX
 PS Claim 2; Fig 7; 184pp; Japanese.
 XX

CC The invention relates to the isolation of novel G-protein coupled
 CC receptor (GPCR) genes and their encoded proteins. This sequence
 CC represents the protein sequence of a mouse GPCR. The DNAs and proteins
 CC are used for the treatment, prevention and diagnosis of disorders of
 CC the nervous system. The proteins and its fragments are also promoters
 CC of somatostatin excretion.
 XX
 SQ Sequence 188 AA;

Query Match 50.8%; Score 517; DB 21; Length 188;
 Best Local Similarity 57.1%; Pred. No. 1e-49; Length 188;
 Matches 113; Conservative 22; Mismatches 51; Indels 12; Gaps 4;

Qy 1 MEIISKLFILLTATSSLLTSNIFCADELVMSNLHRSKENDKYSEPRGYPKG--ERSLN 58
 Db 1 MEIISLKRFIILLTATSSFLTSNTFTDEFMMPHPSKSGDGKYSQLRGIPKGVKERSVS 60

Qy 59 FEELKDWGPKNVKIMSTPAVNKMPHSFANLPLRFGNNVOERSAGATANLPLRSGRNNEV 118
 Db 61 FOELKDWGAKKIDIKMSPAPANKVPHSAANLPLRFGNIEDRRSPARA-----NNEA 112
 Qy 119 SLVRVPLNLPORFGRTTAKSVCRMLSDLCOGSMHSPCANDLFYSMTCHQHOEIQNPQOK 178
 Db 113 GTSMHFFSLPQRFGR-TTARRITKTLAGLPQKSLHSLASSESLYAMTROHOEIQSPQEK 170
 Qy 179 SRRLFKKIDDAELKQEK 196
 Db 171 TRGAFVETDDAERKPEK 188

RESULT 14
 AAG67751
 ID AAG67751 standard; Protein; 188 AA.
 AC AAG67751;
 XX
 DT 10-DEC-2001 (first entry)
 DE
 DE Amino acid sequence of a murine prolactin secretion regulating agent.
 KW Human; prolactin secretion; hypocoovarianism; sperm development;
 KW osteoporosis; lactation disorder; hypothyroidism; kidney failure;
 KW menopause; hyperprolactinemia; pituitary tumour; dienecephalon tumour;
 KW menstrual disorder; stress; autoimmune disease; prolactinoma; sterility;
 KW impotence; amenorrhea; lactorrhea; hyperpituitarism; Sheehan's syndrome
 KW Chauri-Frommel syndrome; Argonz-del Castillo syndrome; lymphoma;
 KW Forbes-Albright syndrome; spermatogenesis disorder.
 XX Mus sp.
 XX WO200166134-A1.
 XX 13-SEP-2001.
 XX 06-MAR-2001; 2001WO-JP01716.
 PR 06-MAR-2000; 2000JP-0065752.
 PR 07-DEC-2000; 2000JP-0378001.
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX Matsumoto Y, Watanabe T, Hinuma S, Habata Y, Yoshida H;
 PI WPI; 2001-596812/67.
 DR N-PSDB; AAH78503.
 XX New polypeptide, useful as a vasotropic, tranquilizing,
 PT immunosuppressive and gynecological agent comprises the prolactin
 PT secretion regulator
 XX
 PS Claim 2; Page 156-157; 180pp; Japanese.

CC The present sequence represents a murine polypeptide which is a prolactin
 CC secretion regulating agent. The prolactin secretion regulating agent
 CC polypeptide and polynucleotide are used for the treatment and prevention
 CC of hypocoovarianism, sperm under development, osteoporosis, menopause,
 CC lactation disorders, hypothyroidism, kidney failure, hyperprolactinemia,
 CC pituitary tumour, dienecephalon tumour, menstrual disorders, stress,
 CC autoimmune disease, prolactinoma, sterility, impotence, amenorrhea,
 CC lactorrhea, hyperpituitarism, Chauri-Frommel syndrome, Argonz-del
 CC Castillo syndrome, Forbes-Albright syndrome, lymphoma, Sheehan's syndrome
 CC and spermatogenesis disorders.
 XX
 SQ Sequence 188 AA;

Query Match 50.8%; Score 517; DB 22; Length 188;
 Best Local Similarity 57.1%; Pred. No. 1e-49;
 Matches 113; Conservative 22; Mismatches 51; Indels 12; Gaps 4;

[illegible]

```
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158212.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 9.6%; Score 98; DB 21; Length 354;
Best Local Similarity 29.3%; Pred. No. 0.036;
Matches 53; Conservative 20; Mismatches 66; Indels 42; Gaps 11;

QY 2 EIISSKLLIITLATSSLLTSNIFCADELVMNSLHNSKENYDKYSEP---RCYPKGERSLN 58
Db 13 OLISAKEDIKVLRLTK-----FCHPILVRLGWHDACTYNNKEEWPRLGGANG--SLR 63
QY 59 FE-ELK---DWGPKNVIKMSTPAYNKMPH-SFANLPLRFRGNVOERSAGATANLPLRSG 113
Db 64 FEELKHAANAGLLNALKLIQPLKDKYPNISIYADL-----FOLASATA-IEEAGG 112
QY 114 RNMEVSLVRRVNPQLPQRFGR-TTAKSVCRMLSDLCQGSWHSPC--ANDLEYSMTQHOE 170
Db 113 -----PDIPMKYGRVDVVAPEQCPEEGRLPDAGPSPADHLRDFYRMGLDDKE 161
QY 171 I 171
Db 162 I 162

Search completed: March 26, 2003, 10:13:55
Job time : 36 secs
```